UNIVERSITY OF HAWAIʻI AT MĀNOA

DOCTORAL DISSERTATION

A Differential Study of Nucleosynthesis in Open Star Clusters

Author:
Michael Gavin LUM

Supervisor:
Dr. Ann Merchant BOESGAARD

Dissertation Committee:
Fabio Bresolin,
Rolf-Peter Kudritzki,
Lipyeow Lim (University Representative),
Roberto Mendez

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Astronomy

in the

Institute for Astronomy
Department of Physics and Astronomy

December 5, 2018

UNIVERSITY OF HAWAI'I AT MĀNOA

Abstract

A Differential Study of Nucleosynthesis in Open Star Clusters

by Michael Gavin LUM

Measuring the chemical composition of stars in open clusters provides the best information on the origin and evolution of these stars and clusters. When we compare compositions of the main sequence stars with those of the evolved (giant) stars in the same cluster, we can measure the results of hundreds of millions of years of fusion processes in the cores of these stars.

This work presents an extensive elemental catalog of 24 individual elements, 7 in multiple ionization states, measured from high-resolution spectra ($R > 40,000$) of 230 stars in 8 open star clusters. We perform a differential analysis between the main sequence and giant stars to measure elemental evolution from core nucleosynthesis. Using the main sequence turn off age of each cluster, we also analyze the temporal/mass effects on the elemental evolution.

We also present rudimentary artificial intelligence and machine learning techniques used to analyze the spectra, determine cluster membership, and derive stellar atmospheric parameters and elemental abundances

Acknowledgements

Primarily, I wish to thank my advisor, Ann Boesgaard, who paitently mentored me in the arcane arts of spectral analysis, who allowed me to pursue outreach and educational opportunities while I worked on this project, and who patiently tolerated my journey (-ies) down the "rabbit holes" of Bayesian Analysis and Machine Learning.

I would like to thank my committee members Fabio Bresolin, Rolf-Peter Kudritzki, Lipyeow Lim, and Roberto Mendez, and for their questions, feedback and suggestions.

I wish to thank the photographers who allowed me to use their work to bring life to my clusters: Giuseppe Donatiello, Jerry Lodriguss, Greg Parker, Jim Thommes, and Alson Wong. Specific credits accompany the images.

This research has made use of the Keck Observatory Archive (KOA), which is operated by the W. M. Keck Observatory and the NASA Exoplanet Science Institute (NExScI), under contract with the National Aeronautics and Space Administration.

I am extremely grateful to the principal investigators (PIs) of the original observations of the archived data used in our research: J. Cohen, B.Jones, G. Marcy, and S. Schuler, along with many others whose names do not appear in the archival data. I also wish to thank the maintainers of the digital online archives for their respective instruments, and the observatories and staff for the archives and instruments. When available, I include any requested attributions, below.

This work has made use of the VALD database, operated at Uppsala University, the Institute of Astronomy RAS in Moscow, and the University of Vienna.

This publication makes use of data products from the Two Micron All Sky Survey, which is a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center/California Institute of Technology, funded by the National Aeronautics and Space Administration and the National Science Foundation.

I acknowledge the ESO's HARPS and UVES staff and archives for providing our solar calibration spectra.

The National Solar Observatory provided one of the solar baseline spectra used herein.

Contents

List of Figures

List of Tables

To:
Audrey, who has been my support throughout my journey, and who holds my heart.
My students, past, present, and future, whose thirst for knowledge I hope to fulfill.
The children of Hawai'i, both young and old, whose joy in seeing the sky for the first (or fiftieth) time, still inspires me.
...and most especially, to Paul, who opened my eyes to seeing the sky in a whole new way.

Introduction

> Near [Perseus'] left thigh move the Pleiades, all in a cluster, but small is the space that holds them and singly they dimly shine. Seven are they in the songs of men, albeit only six are visible to the eyes. Yet not a star, I ween, has perished from the sky unmarked since the earliest memory of man, but even so the tale is told. Those seven are called by name Halcyon, Merope, Celaeno, Electra, Sterope, Taygete, and queenly Maia. Small and dim are they all alike, but widely famed they wheel in heaven at morn and eventide, by the will of Zeus, who bade them tell of the beginning of Summer and Winter and of the coming of the ploughing-time.
>
> Aratus of Soli, *Phaenomena*
> c.3rd century B.C.E.

0.1 Star Clusters

I have noticed an increasing level of disappointment when aspiring amateur astronomers and casual stargazers have their first experience with a telescope. Ironically, thanks to incredible images from the Hubble Space Telescope, other ground-based observatories, and even amateur astro-photographers, these astronomy fans have high expectations when they first put their eye to an eyepiece. Their enthusiasm seems to evaporate when they are greeted with a faint, fuzzy blob which the professional astronomer claims is the same as that image they saw online. Explanations involving CCD vs. retinal sensitivity, and stacked, long duration exposures only serve to produce glazed expressions. For those of us who remember our first view through a telescope as one of the first steps on our path to professional astronomy, this visible disappointment is disheartening.

Happily, the (seemingly universal) Human fascination with astronomy still exists with the general public. To keep that magic active, we have to change how we present the sky. Instead of starting with the images (and expectations) which are products of billions of dollars worth of equipment, we simply need to encourage our audience to look up, and find something "different." — a fuzzy spot, a point of light which is brighter, or one with a noticeable color. To start simply, with what we can see. Only after achieving the first step of unaided eye viewing do we point a telescope at it — enhancing what we saw with the prior step. Only after the

learner has seen the object with their eyes and a small telescope, should we introduce the next step of the professional image. It's more interesting to build up excitement, than it is to start with the "answer" from the back of the book. In this way, each additional step adds to the viewer's understanding, slowly building their excitement.

The scientific study of star clusters follows a similar path, from the nebulous, night sky objects, seen by ancient observers, resolved into the groups of hundreds of stars in Galileo's telescope, to the thousands of co-moving objects in Herschel's refractor, into the astronomical laboratories we see today. Each new technology step adding to our excitement and understanding.

This project contains an extensive material and nucleosynthetic analysis of our target stars, representative of our field's cutting edge. Essentially, jumping straight to the latest and greatest "back-of-the-book" textbook answer. However, this project also presents the "next step" , advancing our understanding of star clusters through new digital-based analysis techniques, specifically automated, (artificially) intelligent analysis, and machine learning algorithms.

0.2 Spectra, Archives, and Automation

My primary goal for this project is to produce a complete elemental abundance catalog of 8 clusters. I expect that the resulting data will appeal to many astronomers within the stellar, galactic, and planetary sub-fields, along with others which I cannot foresee. While I believe this is, in and of itself, an admirable goal, I embarked on this project with a second goal in mind — to use data obtained exclusively from digital archives.

Starting in the mid-1970s, telescope observations have used CCD technology and the associated digital data storage devices for astronomical observation. The resulting digital data is not limited in a physical sense, as it is with the glass plates, and celluloid negatives of photographic astronomy. In fact, as with the well-known *Moore's Law* (Moore 1965) for semiconductor hardware, we can expect the sum total of the available digital astronomical to grow exponentially. The "Big Data" industry expects the quantity of digital information to double every two years (or less), with machine- and sensor-generated data, like astronomical observations, outpacing more traditional Human- and business-generated data by 5x or more (`https://insidebigdata.com/2017/02/16/the-exponential-growth-of-data`).

As early adopters of digital formats, the astronomical community has borne witness to both the benefits and growing pains of this now not-so-new technology. We have several archives which have now accumulated over three decades of observational data from their respective instruments. Massive sky surveys such as Pan-STARRS (Chambers *et al.* 2016; Flewelling *et al.* 2016) place terabytes of data into their archives on a nightly basis. Our trade publications request that data used for published science either be from "open" data repositories, or be released in such a manner after a short period of exclusivity. Organizations like the Smithsonian Astrophysical Observatory have even created digital versions of their accumulated photographic data.

A simple back-of-the-envelope calculation involving the data used for this project reveals one of the potential benefits of using digital archive data for major projects. The ~250 stellar spectra used in this project were compiled from over 800 individual observations, mostly using 8-10m class telescopes. With a total observation time of over 200 *hours*. With a single night on Keck I costing approximately $53,700, in 2012 dollars (`http://ast.noao.edu/system/tsip/more-info/time-calc-keck`), had this project relied on new observations, the cost would have been well over $1M.

Of course, with great amounts and variety of data comes a similar myriad of issues. While most modern observational facilities record, and offer their data for public use (usually after an "exclusive" period for the original observer), there is no central repository, or archive. Data is archived by individual institutions, and with the lack of a central archive comes the related lack of common data formats and standards. Similarly, data quality control can vary from none, with useless data interspersed with good, to extreme, where potentially valuable data is needlessly excluded (or thrown out).

Over time, the digital growing pains will work themselves out. The *.fits* file format has generally been accepted as the industry norm for astronomical images, although the large number of .fits sub-types still inhibits a complete standardization. Large organizations like the European Southern Observatory and the Space Telescope Science Institute are beginning to offer a common repository for observational data.

When I began this project, I felt that the state of astronomical data archives were ready to support a large-scale (digital) observational project such as this one. I had no illusions that it would be simple - early adopters never have it easy - but it is my hope that the techniques and software I have developed will help to enable future projects involving "Astronomically Big Data".

Addressing large data sets will also require new methods of data processing. Absorption line spectral analysis, the primary technique used in this research, has traditionally been a labor-intensive endeavor. The scientist finds the appropriate spectra region, and manually fits an appropriate line profile to the spectra points. At best, an experienced scientist can fit a single absorption feature once every ten seconds or so. For a few dozen lines in a handful of spectra, this could take an hour or two - barring mistakes and re-measurements. Increase the desired measurements to my line list of 2650 lines, and 250 stars, and suddenly a scientist is looking at 5 weeks of 8-hour days, doing nothing but fitting line profiles.

Anecdotal evidence has led me to conclude that the line measurement process will benefit from an unbiased measurement process. More than a few times, while measuring absorption features manually, I found myself thinking "I just measured another <insert element name> feature, so this one should also fit similarly." A truly unbiased process eliminates both the subjective element of the fitting process (i.e.: the "best fit by eye" element), and that (unconscious) desire for similar features to match.

0.3 Project Objectives

Unfortunately, a myriad of advanced Artificial Intelligence or complex Machine Learning techniques are useless without a science goal. For this project, I combine computational methods, with observational analysis techniques, and (nuclear) physics to examine the compositional evolution of stellar atmospheres over the lifetime of those stars.

Elemental abundance analysis through spectral feature analysis is not even remotely new. The earliest spectral analysis came soon after the first astrophotography, in the mid 19th century. However, only in the last 20 years has the technology been available to measure, with sufficient precision, the abundances of the less-common elements, which yield insight into the processes which occur in the core of a star during its main-sequence lifetime.

This project takes advantage of the serendipitous temporal intersection of the quality of high-resolution digital spectroscopy, and new artificial intelligence and machine learning techniques to produce both new membership rosters and extensive elemental inventories of eight solar metallicity star clusters.

0.4 Document Summary

This document is organized into seven sections. Chapter 1 summarizes historical star cluster research, and gives an overview of similar cluster abundance measurement projects. Chapter 2 goes into more detail on the eight clusters I have chosen to analyze for this project. Chapter 3 describes the machine-learning-based technique I created to determine individual stars' cluster membership, based on a large, public data set. Chapter 4 is an in-depth description of the remaining software, scripts, database(s), and third-party applications I use for analysis and results presentation. Chapter 5 is a case-study and process walk-through for an example cluster, NGC 752. Here I present the abundance measurements for NGC 752 resulting from my analysis. Chapter 6 is a shorter summary of the same process and results for the remaining seven clusters in my target list. The final chapter presents the cross-correlation(s) between clusters, using each cluster as representative of the material composition of a solar-metallicity cluster at that age.

Chapter 1

Background and Historic Reference

1.1 Open Star Cluster Studies

Early astronomers were fascinated with astronomical objects which were "different" — stars which wandered amongst the constellations (*Planetes*) or with "long hair" (*Kometes*). In his "Almagest", a catalog of visible stars and constellations, Ptolemy also described three "nebulous", objects which we now know as star clusters: The Double Cluster (NGCs 869 and 884), M7, and Praesepe. One of Galileo's first telescope observations revealed the nebulosity of the "Beehive" cluster to be comprised of dozens of individual stars, too faint for our unaided eye to discern. This discovery led him to (correctly) postulate that the milky "haze" of our galactic plane to be of the same origin (Galilei 1610).

(A) M44 - Galileo Galilei (B) M44 - Giuseppe Donatiello

FIGURE 1.1: The Beehive Cluster (M44) from Galileo's *Sidereus Nuncius* (Galilei 1610) (left) and a composite image of Sloane Digital Sky Survey data (right)

Through the 18th and 19th century, the dynamics and origin of star clusters provided more material for speculation. William Hershel posited that clusters of stars were the result of "condensation" from their original placement in space. Herschel also was the first to attempt to use these groups of stars to produce a quantitative analysis of our galaxy. In Herschel (1818) he produced a relation between the telescopic magnification power required to resolve star clusters into individual stars, and the distance to those clusters.

Until the advent of astronomy using photographic plates, several decades later, the state-of-the-science for cluster studies lay mainly in their discovery, cataloging, and mapping. Initial photographic efforts, in the latter decades of the 19th century, were dedicated to more precise mapping of individual cluster members. Figure 1.2 is an image of M67 from Isaac Roberts' catalog of star clusters and "nebulae" (including those we now know as galaxies). As decades worth of photographic plates accumulated, astronomers could measure the relative changes of cluster member positions over time — known as proper motion — yielding a quantitative measure for determining cluster membership. Kretz (1900) was an early study which showed that stars in the nearby open cluster in Coma Berenices were co-moving. In a series of publications in the 1910s and 1920s (e.g.: Shapley (1917) (part I) - Shapley & Richmond (1921) (part XVII)), Harland Shapley, using the Mount Wilson Observatory, produced an extensive study of photometry in a variety of star clusters. He included studies of both globular and open clusters, as well as the two known "co-moving groups" known at the time.

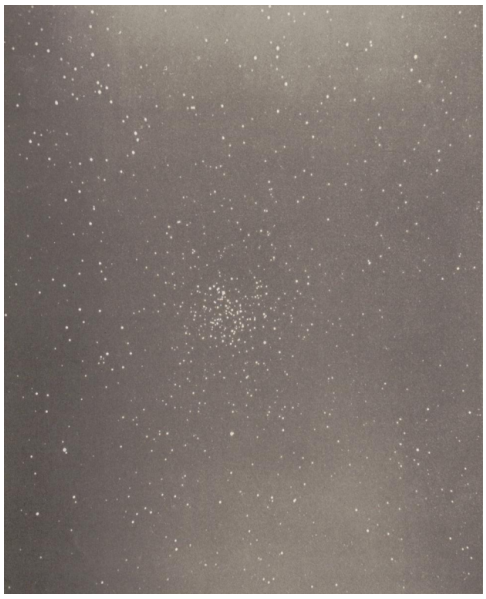


FIGURE 1.2: M67, from Roberts (1893).

The influx of new photometric catalogs, combined with the stellar classification work of Maury & Pickering (1897) and Cannon & Pickering (1901) allowed Russell (1914) to notice that stars lay upon a sequence, defined by their luminosity, and their classification (later temperature). Figure 1.3 is an early version of the *Hertsprung-Russell* (H-R), or Color-Magnitude (CMD) Diagram. Stellar classification (later temperature) is plotted on the horizontal axis, with luminosity (or actual magnitude, or mass) on the vertical. Most importantly, by classifying stars by their spectral features, the new H-R diagrams revealed that there were two distinct populations of the "cooler" stars. The more luminous population were called "giants", an appropriate term which remains today.

The relatively simple act of classifying stars by their spectral type and luminosity led to the next step in cluster analysis. Based on the observation that stars of a particular type and luminosity occupied the same location on the diagram, the distance to a given star could be estimated by scaling its apparent to its observed luminosity. When it was observed that stars from a single cluster also fell on a similar diagram, a much more accurate distance gauge was derived by scaling the "main-sequence" of the cluster to the more general diagram (eg: Johnson & Knuckles 1955). The combination of main-sequence fitting with the earlier discovery of Cepheid variable periodicy as a function of luminosity (Leavitt 1908) refined the precision of distance measures to nearby galaxies.

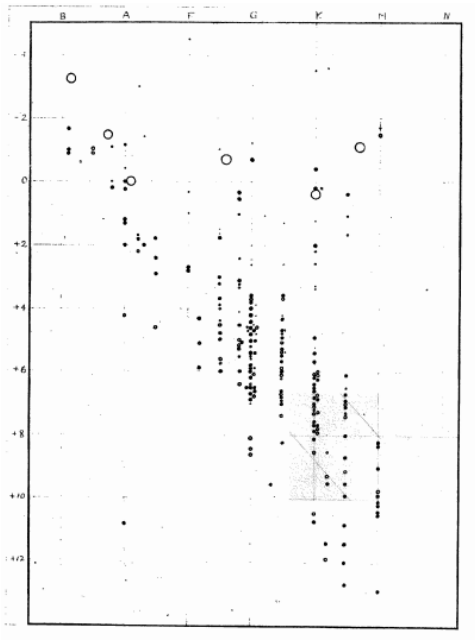


FIGURE 1.3: An early spectral type - luminosity diagram, from Russell (1914).

Early observers believed that a given star's location on the H-R diagram was a function of its evolutionary phase. New, young, hot, luminous stars began their life at the upper end, slowly cooling over time, losing energy to gravitational collapse, to end up at the cool, red, dim tail. It took a greater understanding of a stars' nuclear fusion process, and better modeling of the hydrostatic conditions during a star's lifetime (See: Chandrasekhar 1939, for an early review) to clarify that a star's location on the main sequence was fixed for the majority of its lifetime and was a combination of its initial mass, and the nature of the fusion processes in its core.

Our modern understanding of star clusters attributes a common origin for all members of a cluster from a single interstellar molecular cloud (review paper by Shu, Adams & Lizano (1987)). The massive cloud, containing the ashes of previous generations of stars, is homogeneously mixed, thus as it fractures into protostellar cores, each resulting star starts with the same homogeneous material composition. In an observational study of 40 Milky Way Giant Molecular Clouds (GMCs), Murray (2011) found their lifetime, from initial collapse to dispersal by radiation pressure from massive stars, to be on the order of 25My. Comparing that timescale to the age of the youngest cluster in this study, the Hyades at 750My, there is no need to differentiate between stars potentially formed at 725Mya from those at 775Mya. Essentially, while the mechanism and exact timescale of molecular cloud collapse is still an active field of research (review paper: Kennicutt & Evans 2012), for the intent of this work, we can assume that the stars in our target clusters were formed simultaneously, and of identical initial composition.

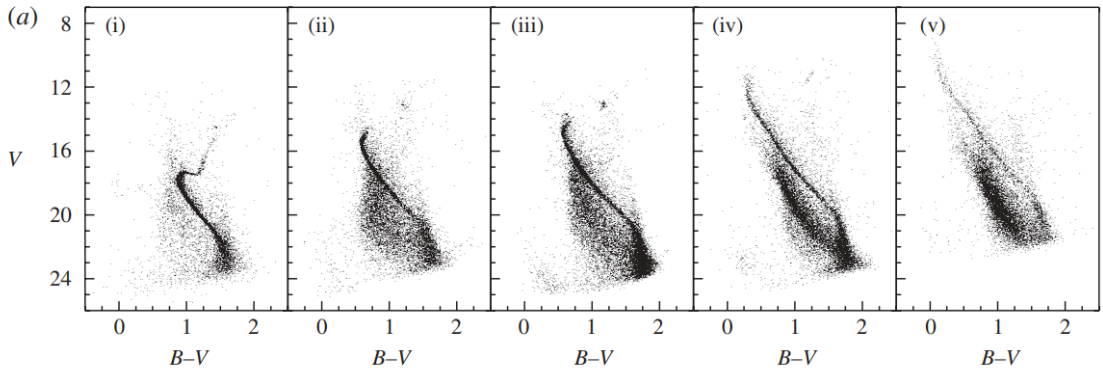


FIGURE 1.4: Figure 2a from Kalirai & Richer (2010): Color-Magnitude Diagrams (CMDs) for five open clusters, in order of decreasing age, from left to right. (i): NGC 6791 (10GY), (ii): NGC 6819 (2.5GY), (iii): NGC 7789 (1.7GY), (iv): NGC 2099 (500MY), and (v) NGC 2168 (200MY). Photometry from the CFHT Open Star Cluster Survey (Kalirai *et al.* 2001).

Setting aside the historical review of star cluster studies, it is important to discuss the implications of the homogeneous and co-evolutionary nature of cluster members. When cluster stars are placed on a Color-Magnitude diagram, cluster members lie along a clearly defined main-sequence, where the star's luminosity is a function of its "color" (Figure 1.4). The correlation is similar to that seen on CMDs made of random "field" stars. Alternatively, since stellar luminosity is proportional to its mass (e.g.: Kuiper 1938), and spectral type to (effective) surface temperature (e.g.: Habets & Heintze 1981), the main sequence can also represent a correlation between mass and temperature. However, as shown in Figure 1.4, while all clusters have a similar main-sequence, they differ at the point where the highest mass stars leave the main sequence. This Main-Sequence Turn-Off (MSTO), can be seen as a function of the cluster's age. Alan Sandage (Sandage 1955) was the first to publish the relationship between cluster age, and the shape and location of the MSTO (Figure 1.5).

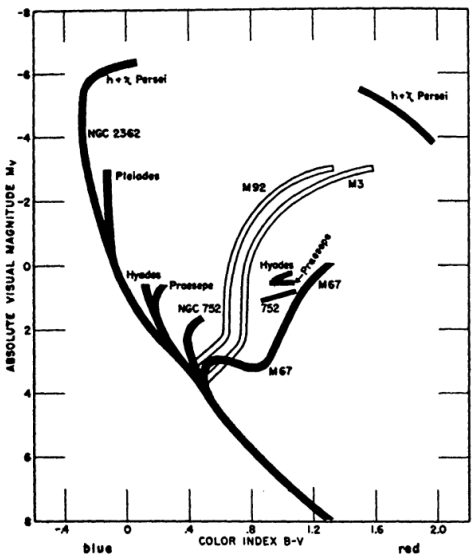


FIGURE 1.5: Figure 1 from Sandage (1955): Cluster Color-Magnitude Diagrams (CMDs) for seven open galactic clusters (solid black lines) and two globular clusters (open lines).

As Sandage observed, stars of the same age (and therefore, cluster stars) trace a single path which can be parameterized by mass. The resulting function is known as an *isochrone*. For this work, I use the exceptional photometry from the GAIA galactic mapping project (Gaia Collaboration *et al.* 2016) to convert photometric, observational data of my target stars into physical parameters. Accurate isochrone creation is an entire field of research, involving stellar interior modeling, convective and radiative energy transfer calculation and simulation, molecular and atomic diffusion, mixing length theory, and many other advanced physical topics. Detailed discussion of the intricacies of the different models is beyond the scope of this work. However, the end results of the theoretical work are often compiled into databases, or "model grids", which I incorporate for my use. Several of these model grids are available online (e.g.: The Dartmouth Stellar Evolution Database Dotter *et al.* 2008). I have chosen to use the PARSEC (Marigo *et al.* 2017) isochrones, mainly due to this model having GAIA G, BP, and RP bands, and for the convenience of the (scriptable) online *CMD 3.0* interface at: `http://stev.oapd.inaf.it/cgi-bin/cmd`. The end product, when combined with my refined cluster membership determination, detailed in Chapter 3, produces CMD/isochrone plots, similar to Figure 1.6.

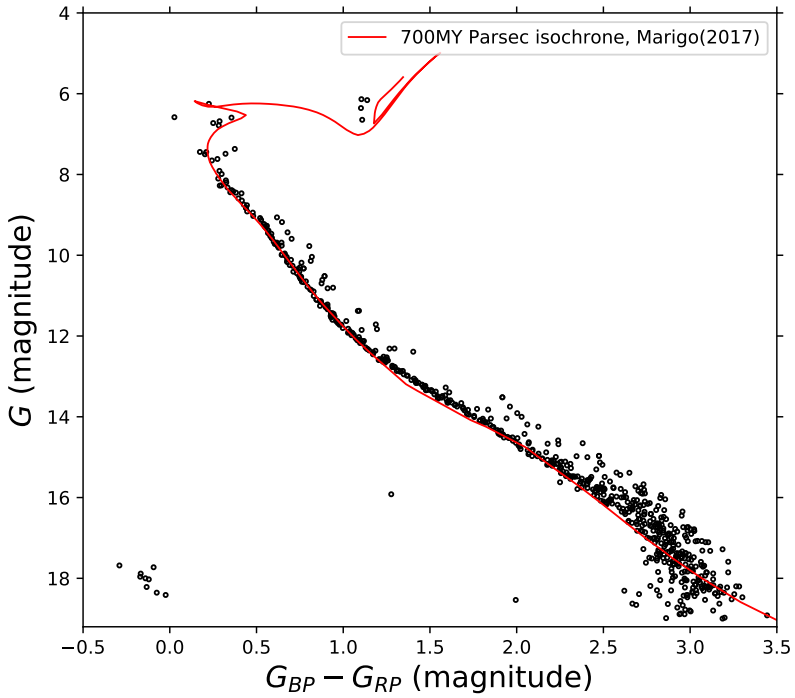


FIGURE 1.6: Marigo *et al.* (2017) isochrone, plotted over GAIA DR2 (Gaia Collaboration *et al.* 2018) photometry, for the Praesepe cluster. When compared to the CMDs in Figure 1.4, this plot has almost no foreground/background stars plotted below and to the left of the main sequence. Membership data from Lum (2018, in prep.)

1.2 Galactic Open Clusters

While the main aims of this work are establishing membership and composition of our target clusters, I feel that it is important to acknowledge the import of other applications of cluster research, not covered here. Specifically, applications which correlate the conditions, as measured within a cluster, to those of the galaxy at the cluster's position. Admittedly, as with a discussion on theoretical isochrone and stellar interior modeling, this topic is far broader than should be covered in this introductory review, so I will restrain the applications discussed here to a handful of examples.

On galactic scales, star clusters are large groups of gravitationally bound stars. As such, they are unlikely to migrate very far in radial distance to galactic center from their origin. When they are compared to other clusters in different regions of the galaxy, each cluster acts as a measure of conditions in its respective region. Additionally, the combined luminosity of cluster members allows them to be observed at greater distances than single stars, and can be used to map the structure of our galaxy (Figure 1.7) (Russeil 2003). Measurements of open star cluster metallicity (and other elemental composition) are used as proxies for the conditions at a given galacto-centric distance, at the time of the cluster formation. Open cluster studies, like those of Friel & Janes (1993) and Friel (1995), have established a galactic metallicity gradient

of about -0.06 dex kpc^{-1} (Chen, Hou & Wang 2003). Using open cluster measurements, Yong, Carney & Friel (2012), and Netopil & Paunzen (2013) noted that the galactic metallicity gradient "flattens" at distances of greater than 12kpc (Figure 1.8).

The same techniques used to map structure and composition of our galaxy can be used on nearby galaxies. For example: clusters in the Large Magellanic Cloud can be sufficiently resolved to determine structural parameters (Werchan & Zaritsky 2011). Cluster ages in the Small Magellanic Cloud can produce a star formation history of that galaxy (Rafelski & Zaritsky 2005). The structure of more distant galaxies, like M101 and M33 can be mapped by measuring the brightest stars in newly-formed clusters (OB associations) (e.g.: Bresolin, Kennicutt & Stetson 1996).

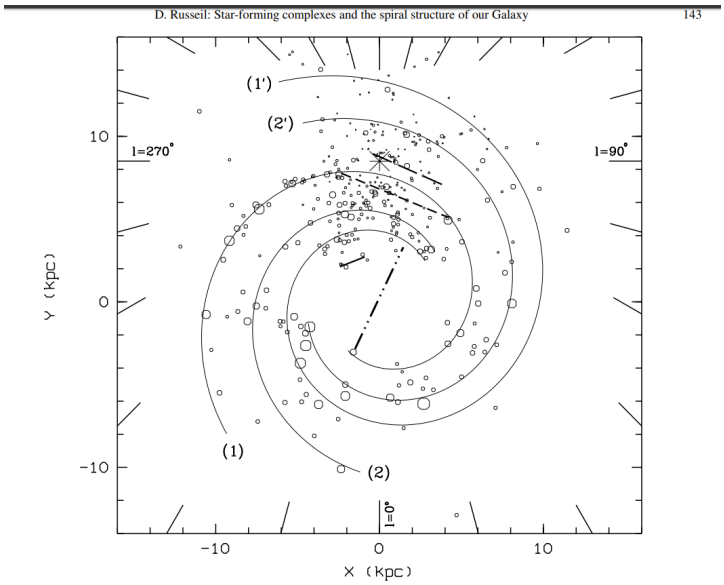


FIGURE 1.7: Figure 5 from Russeil (2003): A map of the spiral structure of the Milky Way galaxy from star cluster and star-forming region observations.

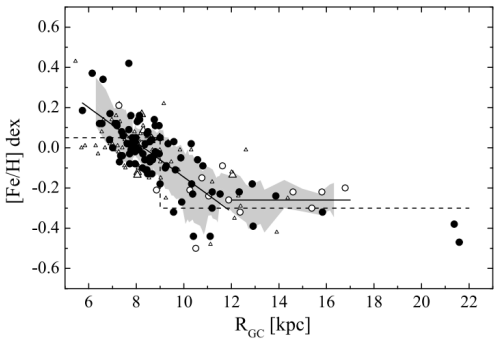


FIGURE 1.8: Figure 4 from Netopil & Paunzen (2013): Open cluster metallicities as a function of galacto-centric radius (R_{GC}).

1.3 The Stellar Life Cycle and Alchemy

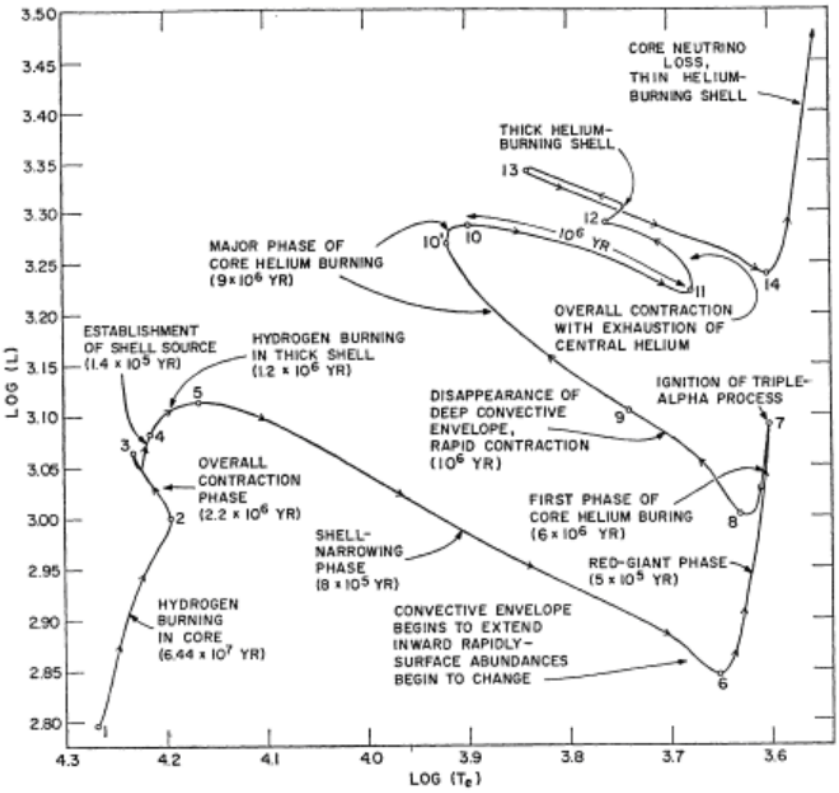


FIG. 1. The path of a metal-rich $5M_\odot$ star in the Hertzsprung-Russell diagram.

FIGURE 1.9: The evolutionary path of a $5M_\odot$ star, fromIben (1967). Note that while the axes are (Log)T_{eff} and (Log) L, as discussed in Section 1.1, they are similar to those in a Color-Magnitude, or HR diagram.

Dwarf stars (interchangeably "dwarfs" or "main-sequence dwarfs") are defined as all stars which lie along the line in the CMD which traces from the faint tail of the main sequence to the MSTO point. In Figure 1.9, Iben plots the evolution of a $5M_\odot$ star's (Log)T_{eff} (effective surface temperature) vs. (Log) L (luminosity). The main-sequence portion of the star's life is labeled as "Hydrogen Burning in Core", and numbered 1-3. Although Iben maps the path of a $5M_\odot$ star, the stages labeled 1-8 are the same for our (approximately) $1M_\odot$ stars. For the purposes of the (somewhat older, >200MY) clusters used in this work, all of our sample stars (both dwarfs and giants) have achieved at least the stable core (H) fusion phase and reside somewhere between the phases Iben has labeled as 1-9. The major difference between our solar mass stars and the more massive star in Iben's diagram is the duration of each phase. Our most massive samples, at about $2.25M_\odot$ can be expected to remain on the main sequence for hundreds of millions of years, at least an order of magnitude longer than the 64MY of Iben's $5M_\odot$ star. We expect to still see our least massive sampled stars, at $0.5M_\odot$ happily fusing H in their cores, for on the order of 50GY.

Once a main-sequence star depletes its core hydrogen fuel supply (regions 3-4 in Figure 1.9), it begins the next phase. With the exhaustion of their core H supply, the post-main sequence stars begin fusing H in a thin "shell" around their core. As the (mostly) inert core collapses, it releases gravitational energy, further heating the shell and atmosphere. These *sub-giants* trace a (nearly) horizontal path along their isochrone, as their atmospheres expand and cool. The increase in energy generation requires that greater portions of the atmosphere become convective in order to transport energy outward (Iben 1991). Eventually, the He core begins to collapse. The additional release of gravitational energy further raises the core/shell temperature, and the star makes an "upward" turn on the CMD.

For solar mass stars, as those in this study, this nearly vertical turn on the CMD marks the onset of the red giant phase. Prior to the red giant phase, He from the p-p and/or CNO fusion process, along with any other nucleosynthetic products are sequestered in the core by an *entropy barrier* (Iben 1977). At the end of the sub-giant phase, the surface convective layer deepens and begins to draw up those fusion products where they can be measured in the stellar atmosphere. While the convective layer reaches at most 80% of the way to the He core, Dominguez *et al.* (1999) show that, even in the case of a $2M_{\odot}$ star, with the least mixing, there is a 30% increase in the He mass fraction at the surface.

As this first "dredge-up" event is occurring, He "ash" continues accumulating in the star's core. For stars in our mass range, the core becomes at least partially (electron-) degenerate. Eventually, conditions in the core reach a temperature of $\approx 10^8 K$ and density of $\approx 10^4 gcm^{-3}$, which are sufficient to initiate He ("triple-α") fusion. Since the core is (mostly) degenerate, the ignition is nearly simultaneous, throughout the core. This core Helium flash (located at the top of the red giant branch, phase "7" in Figure 1.9), releases massive amounts of energy (on the order of $10^{12} L_{\odot}$ for completely degenerate cores). This sudden influx of energy combined with the rapid change in core composition scatters nucleosynthetic products outward from the core regions into convective regions (Mocak *et al.* 2009). With this convection, we expect to see the full range of nucleosynthetic products mixed and visible in the photosphere.

Later phases of more massive stars (phases 9+ in Figure 1.9) do dredge up more material from the core, further altering a star's surface composition. However, for our stellar sample, the stars either do not reach these phases, or spend such a (relatively) short period in these later phases that we were unable to capture any of our sample in these phases.

The primary goal of this work is to create a comprehensive elemental catalog of the target clusters, using mechanical techniques. However, a secondary goal is to compare the abundances of some of these elements which could be enhanced or depleted by nucleosynthetic processes during the main-sequence life of the stars. To do so, I will compare the measurements of the dwarf and giant populations within the target clusters, and by using comparisons of clusters of varying ages, show observational evidence of these core fusion processes.

Chapter 2

Target Clusters

From our vantage point in the galaxy, we can observe on the order of 2000 open, or galactic clusters (Dias *et al.* 2002). Only about half of these have been studied to the point of identifying individual members, and measuring cluster proper motion and radial velocity — an important step in differentiating actual clusters from random overdensities in galactic field stars. However, studies such as Kharchenko *et al.* (2009) are starting to take advantage of large repositories of survey data, and give us a greater statistical understanding of our nearby cluster population.

For the purposes of this project, I was restricted to measuring stars within clusters which are near enough to have high resolution spectra taken of both dwarf and giant stars, as well as a number of other criteria.

During the process of this project, the GAIA mapping project (Gaia Collaboration *et al.* 2016) has released two sets of preliminary data. The unprecedented precision of the photometric data from even the second preliminary data release (DR2: Gaia Collaboration *et al.* 2018) opens new possibilities in cluster membership determination, one of which is detailed in Chapter 3. For the purposes of this cluster overview chapter, I will be frequently referencing the GAIA team's mapping of 32 open clusters (and 14 globulars) from Gaia Collaboration *et al.* (2018), hereafter DR2CMD.

2.1 Selection Criteria

As one of the main goals of this project was to look for evidence of compositional evolution in the atmospheres of stars by comparing the unaltered atmospheres of dwarf stars with the evolved atmospheres of giants, using high resolution spectral analysis, the obvious criterion for cluster selection was that the cluster contain both giant and dwarf stars, and that high resolution spectra were accessible in the selected spectral archives. I also placed an additional restriction that the cluster have (near-) solar metallicity, as I would be using the solar spectrum as a calibration tool for the absorption line measurements.

As some of the elements and ions, particularly C, N, and O, I wished to measure have very faint absorption features, the spectra needed to have relatively high (> 50) Signal-to-Noise ratios. Similarly, the spectra also had to have high resolution ($R \geq 35000$). Additionally, as rotational broadening would "blur" out these fine features, I had to select stars with relatively low rotational velocities. Since stars "spin-down" with age, and red giant stars are only observed in "middle-aged" clusters ($> 700MY$), the restriction of low rotational velocity and the presence of red giants in the cluster are basically the same criteria, where one does not overly restrict the sample population over the other. These factors, and the need to use publicly available archive data limited my search to just a handful of instruments and archives.

I need to emphasize that my selection criteria created a selection bias, in that my results can only be generalized to middle-aged clusters/stars of solar metallicity, relatively close to the solar neighborhood. As this perfectly describes our own Sun, the bias is not particularly concerning, but should be made clear that my results only apply to such stars.

The clusters which met the following criteria are individually detailed in the subsequent sections.

1. Approximately solar metallicity ($-0.20 < [Fe/H] < +0.50$)
2. Spectra from both giant and dwarf populations
3. Archived High Resolution Spectra
4. High Resolution Spectra from one or more of:
   (a) W.M.Keck I Telescope – HIRES
   (b) ESO Unit Telescope 2 (UT2) – FLAMES
   (c) ESO Unit Telescope 2 (UT2) – UVES
   (d) ESO 3.6m Telescope – HARPS
   (e) Observatory of Haute-Provence 1.9m Telescope - ELODIE
   (f) Observatory of Haute-Provence 1.9m Telescope - SOPHIE
5. High Signal to Noise ($S/N > 50$)
6. Low ($< 20 km\ s^{-1}$) Rotational Velocity

2.2 Praesepe (NGC 2632)

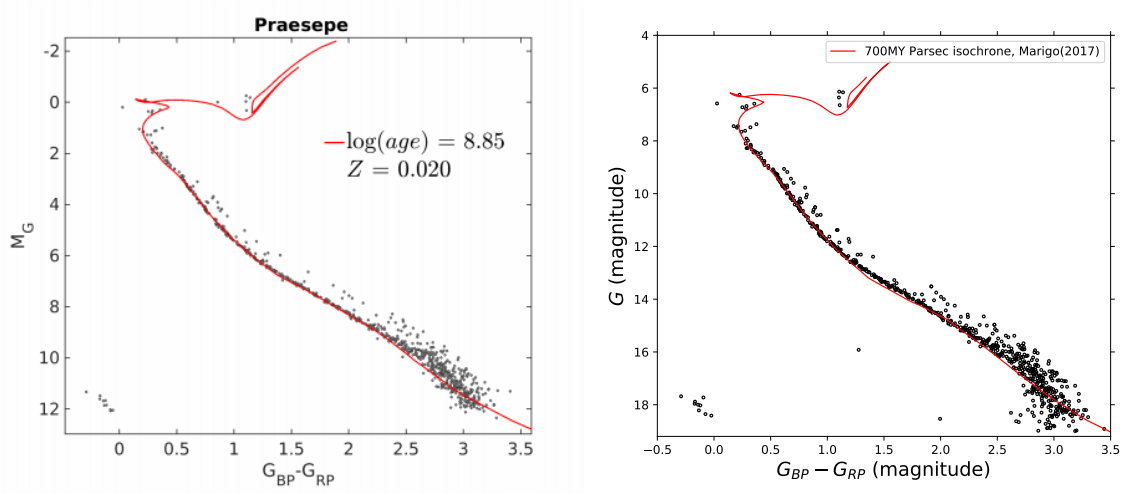
Praesepe has been known since ancient times. Galileo was the first to observe Praesepe in a telescope (Figure 1.1a), and determined that it was a cloud of stars, not the nebulous mass perceived by keen sighted astronomers. The cluster consists of around 800 members (DR2CMD), which fit a 700MY isochrone (Figure 2.2). The 700MY age also corresponds to age estimates derived from main-sequence geometry ($\delta(V), \delta 1$) in Salaris, Weiss & Percival (2004), and the upper limit of white dwarf cooling times in the cluster (Claver *et al.* 2001; Dobbie *et al.* 2004). As can be seen in Figure 2.2, in addition to the single "blue straggler" star on the HR diagram, there is a significant spread among the MSTO stars. A number of them can be better fit with a 600MY isochrone, while the ones which lie red-ward from the 700MY isochrone could be better fit with a 800MY isochrone. Brandt & Huang (2015) attribute the spread to rotational effects, where stars with higher rotation rates can extend their main sequence lifetime. Their age estimate for the cluster corresponds to the older (800MY) isochrone, with more massive, fast rotators fitting to the younger (600-800MY) isochrones. For the purposes of this work, the apparent age difference is not a concern, so I opt to use the age as fit by the 700MY ("median") isochrone.

FIGURE 2.1: Praesepe, M44, or The Beehive cluster. Credit: Digitized Sky Survey (DSS), STScI/AURA, Palomar/Caltech, UKSTU/AAO

The distance modulus for my Praesepe isochrones is 6.35, taken from DR2CMD, which translates to a distance of 186pc, and is similar to the values determined by Loktin (2000) ($DM = 6.16$, 171pc), and Percival, Salaris & Kilkenny (2003) ($DM = 6.22$, 175pc).

Praesepe abundances are relatively well studied. Most recently, Carrera & Pancino (2011) looked at three giants, Yang, Chen & Zhao (2015) examined abundances of 17 elements in A, F, G, and K stars. which they then compared to our work (Boesgaard, Roper & Lum 2013) to determine whether reports of chemical inhomogeneity in the cluster (Vereshchagin & Chupina\inst 2013) could be confirmed. They found no differences in their selected regional groupings across the cluster.

I have selected Praesepe as a target cluster, mainly due to the availability of spectra, but it is also of great interest as a comparison point — essentially a "check" for my automated process. As most studies have found similar (within uncertainties) abundances of their commonly measured elements, I do not expect my process to produce values far from those already produced. Should my process produce statistically different values, I would assume that it was an error in the process, and attempt to "de-bug" the error.



(A) Praesepe HRD - Gaia Collaboration *et al.* (2018)

(B) Praesepe HRD - This Work

FIGURE 2.2: A comparison of the Hertzsprung-Russell Diagrams (HRD) built using GAIA DR2 data, and the Gaia Collaboration *et al.* (2018) (Figure 17) membership (left) compared to the membership technique described in Chapter 3. Isochrone overlay for both is a 700MY Padova isochrone from Marigo *et al.* (2017). Horizontal scales are identical. Vertical scale on the left is corrected for the distance modulus of the Praesepe cluster of 6.35.

2.3 Hyades

FIGURE 2.3: Hyades, with NGC 1647 also visible in the far left. Image by Jerry Lodriguss, AstroPix.com.

The nearest cluster in my study is also probably the most studied. Like Praesepe, the Hyades have been known since ancient times, and are near enough that their membership is well understood and documented. The most recent, in-depth examination of cluster photometry in DR2CMD, places the center of the Hyades at 47.6pc, identical within uncertainty to the value of 46.9pc I obtain from a slightly different method in the next chapter. Traditionally, the distance to the Hyades cluster is small enough that cluster reddening (E(B-V)) is considered negligible — confirmed by the reddening value in DR2CMD, listed as 0.001.

Liu *et al.* (2016) provides the most recent elemental analysis, using 18 elements, and find that the Hyades are actually chemically IN-homogeneous. While this may seem to be a violation of the assumption of chemical homogeneity I make for this project, the differences they measure are well below the uncertainties I have for my different population samples (dwarf and giant). Their small uncertainties are due to two methods which I cannot utilize in my study - first, they restrict their stellar sample to only dwarfs with effective temperatures between 5650K and 6250K. Their narrow stellar type restriction allows them to use a differential analysis, directly comparing the measured equivalent widths with those of a star with well-studied parameters in the same range. Since I am comparing two significantly different stellar types (more, if the F-M stellar types, sub-, and red-giants are considered different), the differential analysis technique (e.g.: Önehag *et al.* 2011), is not practical.

The techniques used by DR2CMD, find 480 members, compared to 251 by Reino *et al.* (2018), using the earlier GAIA DR1 data set (Gaia Collaboration *et al.* 2016), and 197 by Perryman *et al.* (1998). Our membership algorithm finds many more, but see the discussion in the next chapter.

The Hyades are an important part of this study for several reasons. First, along with Praesepe, the Hyades are the youngest cluster (at 700MY) to contain both giants and dwarfs. As such, these two clusters represent the earliest point for abundance comparisons to measure atmospheric evolution over the lifetime of the stars. Secondly, the Hyades and Praesepe are suspected of having formed from the same giant molecular cloud and are members of the "Hyades Supercluster" (Tabernero, Montes & González Hernández 2012). At the very least, due to their nearly identical ages, and highly similar metallicity, having two data points at my youngest cluster age reduces uncertainty.

2.4 IC 4756

IC 4756 is the first of my "middle-aged" clusters. Of course, the middle-aged moniker is purely subjective, as all of these clusters would be considered "old" for an astronomer who studies star-forming regions, or "young" to one who studies globular clusters. For the purposes of this work, Praesepe and the Hyades are considered "young", at $\approx 700MY$. Clusters between 1GY and 2.5GY are the middle aged sample, and M67 and NGC 6791 at 4.0, and about 8.0GY are the "old" open cluster sample. While Netopil *et al.* (2016) and Paunzen & Netopil (2006) place IC 4756 at $< 700MY$, which would place it in the "young" category, I am inclined to trust the age estimate of the DR2CMD work at 990MY. In fact, if I assume that rotational effect are broadening the MSTO, as per Brandt & Huang (2015), a 900MY isochrone represents high rotation stars which have had their main sequence life extended by rotationally-induced mixing, while slower, or non-rotators would fit a 1.2GY isochrone (Figure 2.5).

FIGURE 2.4: IC 4756, Credit: Jim Thommes AstroPhotography, http://www.jthommes.com/Astro/.

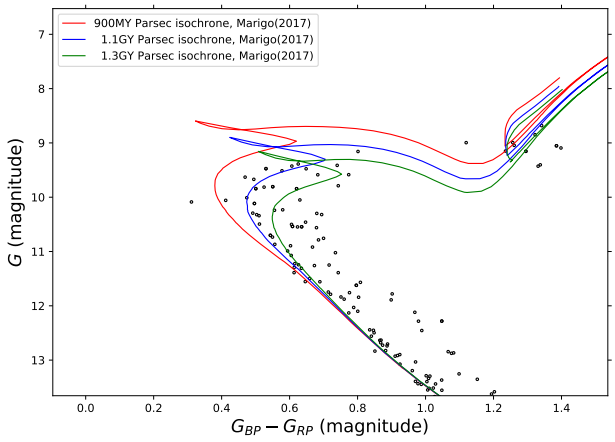


FIGURE 2.5: IC 4756, Main-Sequence-Turn-Off (MSTO) point "close-up", potentially showing the extension of the main-sequence life of 1.2GY stars with high rotation rates, as per Brandt & Huang (2015).

Recent distance measurements place the cluster at 415-483pc (Paunzen & Netopil 2006; Kharchenko *et al.* 2016; Gaia Collaboration *et al.* 2018), corresponding to a Distance Modulus (DM) of 8.4, which was used to fit the isochrones in Figure 2.5. The reddening value in the figure is 0.12 (E(B-V), but applied unaltered as $E(G_{RP} - G_{BP})$, which is suggested by DR2CMD, but is significantly less than that listed by Paunzen & Netopil (2006), at 0.19.

the metallicity of the 500 (DR2CMD) or so ($\approx$ 250, this work) members of IC 4756 has been measured at solar ([Fe/H] = +0.02 Netopil *et al.* (2016)), with abundances of 17 other elements measured by Blanco-Cuaresma & Soubiran (2016) in attempt to establish an elemental "fingerprint" for current and former cluster members

2.5 NGC 752

As with all but one of the clusters in this project, NGC 752 is also a "solar twin" cluster, with a metallicity approximately solar ([Fe/H]= -0.03 ± 0.06 Netopil *et al.* (2016), $+0.04 \pm 0.01$ Blanco-Cuaresma *et al.* (2014), [Fe/H]= -0.063 ± 0.013 Maderak *et al.* (2013), [Fe/H]= $+0.08 \pm 0.04$ Carrera & Pancino (2011)). Carrera & Pancino and Reddy, Giridhar & Lambert (2012) both conducted comprehensive elemental analysis, where the abundances of more than 10 elements were measured, for four 752 giants. There are also multiple studies where a small number of elements, usually Fe plus one or two other, are measured in the cluster (eg: Maderak *et al.* 2013; Böcek Topcu *et al.* 2015). NGC 752 is also one of the clusters Blanco-Cuaresma & Soubiran (2016) attempted to differentiate, using chemical tagging.

FIGURE 2.6: NGC 752, Credit: Jim Thommes AstroPhotography, http://www.jthommes.com/Astro/.

The distance and reddening measures for NGC 752 have been consistent over the past few decades. Distance measures for the cluster haven't varied much, from 400pc in Lyngå (1983) to 450pc in DR2CMD and Kharchenko *et al.* (2016). Reddening values are also fairly consistent, around $E(B-V) = 0.040$ (DR2CMD, Kharchenko *et al.*), with the NASA/IPAC Infrared Science Archive (https://irsa.ipac.caltech.edu/applications/DUST/) showing mean (total) B-V extinction along the line of sight to the cluster of 0.0478 ± 0.0018 (including the "new" corrections from SDSS in Schlafly & Finkbeiner (2011)). As the NASA/IPAC reddening represents the total extinction from galactic material (dust, mostly), along the line of sight to extra-galactic sources, it is best presented as an *upper limit* for the cluster reddening.

NGC 752 is relatively old open cluster, with an age between 1.4 and 1.8 Gyr (1.40Gyr, DR2CMD, 1.45 Gyr Anthony-Twarog *et al.* (2009), 1.6 Gyr Carrera & Pancino (2011), 1.69Gyr Netopil *et al.* (2016), 1.78 Gyr Daniel *et al.* (1994)). Although the DR2CMD research counts 259 stars as members of the cluster, NGC 752 shows evidence of either dynamic or tidal dispersal particularly in the lower main-sequence (Carraro, Monaco & Villanova 2014). While we do not utilize spectra from the late-K and M region affected by this dispersal, the loss of these lower mass stars is possible evidence that NGC 752 is well into the process of disappearing into the galactic field.

For the reasons discussed in Chapter 5, the stars in NGC 752 served as the "proof-of-concept" cluster for the techniques and processes used in this study.

2.6 NGC 3680

FIGURE 2.7: NGC 3680, Wikisky.

Of the four "middle-aged" clusters in this survey, NGC 3680 shows the most main-sequence depletion — a sure sign that it will not "last much longer". Bica *et al.* (2001) declare that 3680 is well along in the process of evolving into a "POCR" (Probable Open Cluster Remnant). Kharchenko *et al.* (2016) only list 16 "high probability" members for this cluster, while Nordstroem, Andersen & Andersen (1997) only list 50 probable and possible members. My analysis in Chapter 3 reveal a few more, but the cluster is definitely very sparse, with significant depletion amongst the less massive member population.

Cluster distance and age estimates are reasonably consistent, with the distance to the cluster in the 1000pc range (Friel *et al.* (2002) : 1070pc, Anthony-Twarog & Twarog (2004) : 1100pc, Kharchenko *et al.* (2016) : 946pc), and the associated B-V extinction values also consistent with each other (Friel *et al.* (2002) : 0.05, Anthony-Twarog & Twarog (2004) : 0.06, Kharchenko *et al.* (2016) : 0.062) and with the upper limit of the NASA/IPAC galactic extinction of 0.0810 ± 0.0015. The cluster age estimates are in the same range as NGC 752 (Friel *et al.* (2002) : 1.5GY to Anthony-Twarog & Twarog (2004) : 1.85GY), making these two clusters a "middle-aged" pair, in the same sense that the Hyades and Praesepe comprise a "young" pair.

The metallicity of NGC 3680 is not as well constrained as NGC 752, and other clusters in this survey. Estimates of [Fe/H] vary from solar (Heiter *et al.* (2014) and Netopil *et al.* (2016) both measure [Fe/H]=-0.01) to significantly lower (Anthony-Twarog & Twarog (2004) : -0.14 ± 0.03). Pasquini, Randich & Pallavicini (2001) claimed a $[Fe/H] = -0.17 \pm 0.12$ only after applying a +0.10 dex "systematic" correction, which may even indicate a lower metallicity. In addition to their measurement of 16 other elemental species, Blanco-Cuaresma *et al.* (2015) measured [Fe/H] in NGC 3680 at -0.10 and -0.13 for their dwarf and giant populations, respectively. Their measurement seems to offer support to the lower than solar metallicity camp.

2.7 IC 4651

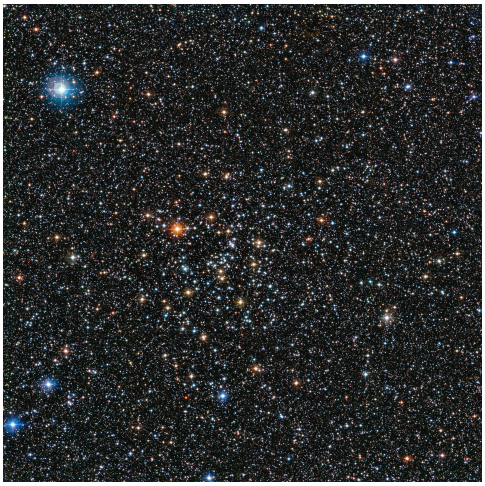


FIGURE 2.8: IC 4651, from the MPG/ESO 2.2-metre telescopes in La Silla, Chile (Credit:ESO).

At about 2.0GY (DR2CMD, or 1.86GY from Netopil *et al.* (2016)), IC 4651 is the oldest of the three "middle aged" clusters. It lies in the galactic plane, just 20° from the galactic center. The NASA/IPAC galactic reddening in this direction is high (0.213), but since IC 4651 is between us and galactic center, interstellar reddening is between 0.040 (DR2CMD) and 0.121 (Kharchenko *et al.* 2016). My best-fit isochrone fit required a reddening of 0.08, a happy median of the two reported values. Interestingly, my dating for this cluster by isochrone fitting (See Chapter 3) places the cluster age upwards of 2.0GY, and closer to 2.5GY. This isochrone was fit using a distance modulus (DM) of 9.72, corresponding to the Dias *et al.* (2002) and Kharchenko *et al.* (2016) cluster distance measurements of 890pc. The DR2CMD values of 2.0GY, 950pc (DM=9.889), and E(B-V)=0.040 produces a questionable isochrone fit (Figure 2.9). Enough so that I could believe that there was an error in transcription of the cluster statistics.

The stars used to model the isochrone in Figure 2.9 come from my membership set of $\approx$ 450 probable members (Chapter 3), using the metallicity from Netopil *et al.* (2016) of $[Fe/H] = +0.12$. The GAIA group did arrive at a larger membership sample of 932 stars, based only on location and proper motion cuts, potentially explaining the difference in our isochrone fits.

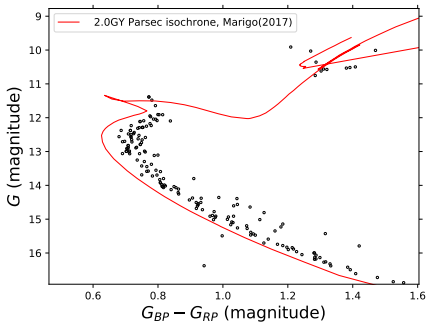


FIGURE 2.9: IC 4651 and isochrone using the GAIA (Gaia Collaboration *et al.* 2018) suggested parameters. The mis-fit, especially near the turn-off may indicate a transcription error in the GAIA cluster paper.

As IC 4651 is very similar in both age and metallicity to M67, Blanco-Cuaresma & Soubiran (2016) attempted to see if abundance measurements from 28 elements could be used to "tag" individual stars as members of individual clusters. While they found that uncertainties in individual element measures, probably caused by unaccounted-for factors such as rotational mixing in atmospheric models, prevented accurate tagging, their comprehensive catalog provides an excellent comparison work for my project.

2.8 M67 (NGC 2682)

FIGURE 2.10: M67 Imaging: Greg Parker, Processing: Noel Carboni

As indicated by its inclusion in Messier's catalog (Messier 1781) as object 67, NGC 2682 is one of the oldest known open clusters. "Oldest known", in both the sense of early documentation, and in terms of age. The stars in M67 are between 3.5 and 4.1GY (Netopil *et al.* (2016) - 3.45GY, DR2CMD - 3.5GY, Meynet, Mermilliod & Maeder (1993) - 4.0GY, and Paunzen & Netopil (2006) - 4.09GY). As with NGC 6791, another old cluster, M67's age and persistence in the face of galactic tidal effects can be attributed to its large population. DR2CMD count 1200 stars as members, while the Montgomery, Marschall & Janes (1993) place the cluster mass at 724$M_{\odot}$. As with my "middle-aged" clusters, M67 is also showing its age, in that it has mass segregation (Montgomery, Marschall & Janes) with the most massive stars concentrated toward the center of the cluster.

M67's physical parameters are well understood, partially because of the low reddening in the line of sight (0.037 - DR2CMD, 0.050 - Kharchenko *et al.* (2016) and Paunzen & Netopil (2006)). The cluster lies at a distance of about 850pc (820 - Paunzen & Netopil (2006), 880 - DR2CMD, 890 - Kharchenko *et al.* (2016)). At a DEC of $+11^{\circ}$, it is also well represented in available spectra from both my Northern and Southern sources.

Elemental abundances in M67 are very well studied, with metallicity determined to be approximately solar (e.g.: Netopil *et al.* 2016, [Fe/H]=+0.03). As noted in the discussion of IC 4651, M67's quasi-twin, Blanco-Cuaresma & Soubiran (2016) analyzed 28 elements to determine if chemical tagging could show the two clusters shared a common origin.

2.9 NGC 6791

NGC 6791 is definitely the most anomalous cluster in this study. Age estimates vary greatly, from the estimate by Bedin *et al.* (2008) of the white dwarf cooling age of 6.0GY, to MSTO dating by Netopil *et al.* (2016) of 7.00GY, Paunzen & Netopil (2006) 7.85GY, Chaboyer, Green & Liebert (1999) 8.0GY, and King *et al.* (2005) 9.0GY. In any case, for a cluster to persist for the multiple billions of year lifetime of NGC 6791, it has to be massive, and/or be located outside of the galactic plane. NGC 6791 appears to fulfill both conditions. Even assuming a reasonable amount of low-mass star depletion, NGC 6791 is massive, with easily over 4000 members (King *et al.* 2005). It also has an unusual "boxy" orbit (Bedin *et al.* 2006), which takes it as much as 1500pc above (and below) the galactic plane. Even with the "protection" of an inclined orbit, and large initial mass, King *et al.* (2005) note mass partitioning, with lower-mass stars preferentially occupying the outer regions of the cluster. Mass partitioning is one of the first steps in cluster dispersal. Low-mass stars in the outer regions of the cluster are more subject to outside influences which separate them from the cluster, further reducing the overall cluster mass, allowing more stars to break free.

FIGURE 2.11: NGC 6791, Credit:Digitized Sky Survey (DSS), STScI/AURA, Palomar/Caltech, UKSTU/AAO

The age and location of 6791 aren't its only anomalous characteristics. The metallicity of the cluster is significantly *super*-solar. Our own earlier work (Boesgaard, Lum & Deliyannis 2015) measured [Fe/H] of $+0.30 \pm 0.02$. Other [Fe/H] measures have ranged from +0.32 (Worthey & Jowett 2003) to +0.42 (Netopil *et al.* 2016) to +0.45 (Anthony-Twarog, Twarog & Mayer 2007). Bedin *et al.* (2006) hypothesize that NGC 6791's exceptionally high metallicity may be a result of forming in the central region of the galaxy, where metallicity is higher (See Figure 1.8). Bedin *et al.* trace the cluster's "boxy" orbit to a galacto-centric passage of $R \approx 3$kpc, which would place it in a region where cluster metallicities are in the +0.40 range (Figure 1.8, again).

Regardless of its origin, or orbit, NGC 6791 lies at a helio-centric distance of about 4500pc (Paunzen & Netopil (2006) : 4418pc, Kharchenko *et al.* (2016) : 4926pc). As it is currently about 1000pc above the galactic plane, interstellar reddening is low (for 4500pc), at 0.117 (Kharchenko *et al.* 2016) to 0.17 (Paunzen & Netopil 2006). Using the original method from Schlegel, Finkbeiner & Davis (1998), the NASA/IPAC galactic reddening for that line of sight was calculated at +0.155, identical to the value in Anthony-Twarog, Twarog & Mayer. However, the corrections from SDSS data revise the NASA estimate down to 0.1330 ± 0.0019. However, based on isochrone fitting, particularly along the red giant branch (Chapter 3), I am inclined to use the higher reddening value of 0.155.

Chapter 3

Cluster Membership Determination

An important first step in my cluster analysis is verifying that a given star (and the associated spectra) is a member of the cluster under study. In previous work (Boesgaard, Roper & Lum 2013; Boesgaard, Lum & Deliyannis 2015), we selected stars from established membership lists (Mermilliod & Mayor 1999; Montgomery, Janes & Phelps 1994, respectively). As I am using spectra from digital archives, where the prior science goal was probably related to the star's associated cluster, I could adopt a similar verification process for this work. However, I believe that having a generic membership determination algorithm anticipates future advances in the availability of both photometric and kinematic data, as well as continually growing spectra datasets (particularly from multi-object instruments, surveys, and wider availability of spectra archives). Additionally, historic membership classification has been based on the technology of the time, which in many cases relies on photographic plates. Higher precision, and more accurate photometry from more technically advanced instruments, like the GAIA satellite mission, should merit re-assessment of cluster membership. For this work, I have implemented a generic Bayesian (statistical) evaluation method, using machine learning techniques, drawing from the second GAIA (Gaia Collaboration *et al.* 2016) data release (DR2) (Gaia Collaboration *et al.* 2018), but equally applicable to other "complete" data sets (which include sky coordinates, proper motion, radial velocity, and parallax).

3.1 Machine Learning Data Sets and Models

The process I have implemented uses a simplified form of *semi-supervised (machine) learning* (See Zhu (2008) for a literature survey/summary). The procedure takes a set of known star cluster[1] members, as a *training set* and creates a series of statistical models. The data points in the training set are labeled as belonging to one or more populations, thus also being referred to as *labeled input*. Individual models are combined to create a common statistical distinction for the known cluster members in the dataset. For astronomers, this process is very similar to selecting a set of priors to combine to create a posterior distribution in Bayesian analysis. I then use this (posterior) model on the full set of GAIA data for the cluster region, to rate all sources with various scores, which I then use to categorize the stars as "cluster member", "possible member" or "possible non-member". By extension, I consider the data points which do not achieve the minimum score for "possible member", and which were not in the original training set, as "non-members".

[1] Unfortunately, the fields of computational data mining and machine learning use the term "cluster" to refer to clustering of data points with similar characteristics. For this work, I equate "cluster" with "star cluster", and will avoid using the term "cluster" for data clustering.

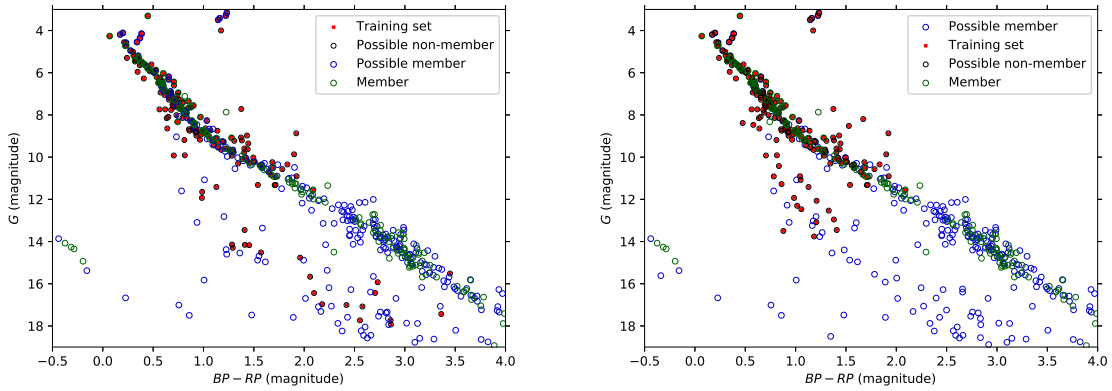
Semi-supervised machine learning relies on iteration, utilizing "self-training" techniques. After using the literature member list, each subsequent iteration uses a new member list, derived from the previous step, as a new training set. While there are a number of convergence analysis methods (several examples in Culp & Michailidis (2007)) to determine when to terminate the iteration ("stopping conditions"), I used a simple subjective analysis of the main sequence manifested in the color-magnitude ($(B-R)vs.G$) diagram as an independent indicator for the iterative process' stopping condition (see 3.2). For nearby clusters ($Parallax > 2.0$) the main sequence is clearly defined after two iterations. For more distant clusters, or clusters in more crowded regions (like M7 in Scorpius), a third iteration is needed.

3.1.1 Training Sets and "Unlabeled Data"

A generative model requires a set of data points with known characteristics, from which it creates probabilistic models to distinguish between groups within the dataset. For each star cluster, I select the initial training set from a known list of cluster members. The literature sources used in selecting these members are shown in Table 3.1. The basic criteria I required to select a given catalog was that the literature contain RA and DEC coordinates, along with and either a binary (member/non-member), or probability-based membership criterion for each member. Since the input for a generative model only requires a binary designation, I converted probability- or score-based membership values into member/non-member binary distinctions using either limits set in the literature, or an arbitrary limit proportional to 60/100 if no literature limit was given. The literature-designated cluster members then became the "positive" labeled data set for the generative model.

In four cases, I selected my membership criteria based on the "Member" classification in the Simbad database Wenger *et al.* (2000). In these cases, while literature-based membership classification existed, the stellar coordinates were based on an X-Y system from photographic images. While translating the photographic coordinates into RA and DEC coordinates is possible, I felt it acceptable to use the coordinate translation and, therefore, the membership scores as listed in the Simbad database.

Once I have a list of members, I then match each star to it's nearest (spatial) match in the GAIA catalog. The spatial matching is performed by the Astropy (Astropy Collaboration *et al.* 2013) SkyCoord "match_catalog_sky" function. All further analysis is performed using the spatial, photometric, and kinematic as reported in the GAIA DR2 database for a given star. In most cases, the literature sources are more limited in depth (in the nearest equivalent to the GAIA G band) when compared to the GAIA survey. Since it is possible that the GAIA survey has cataloged a dimmer source, unseen in the literature survey, which lies closer to the specified coordinate, I add additional criteria to the matching algorithm and try to match the reported (photographic, B, or V) magnitude to the GAIA G magnitude. Since I do not have an exact conversion from the existing systems to GAIA's G-band, I allow the algorithm to match as long as the GAIA and reported magnitude values are within 1.0. Figure 3.1 places the Perryman *et al.* (1998) and GAIA DR2 matched sources (red) on a color-magnitude diagram. The Perryman sources were taken from the Hipparcos (Perryman *et al.* 1997) survey, and thus are limited to sources brighter than V=12.0. The left pane shows my algorithm's matches, using only positional matches. The right pane shows the combination of position and magnitude matches. Note how the Perryman sources, previously matched to faint (GAIA G > 14.0) sources, are now paired to GAIA sources with closer brightness matches.



(A) (B)

FIGURE 3.1: Side-by-side comparison of the Color-Magnitude Diagram of Hyades. Figure 3.1a (left) contains the initial (Perryman *et al.* 1998) member list in red, matched to the nearest (in proximity) GAIA DR2 source. Figure 3.1b (right) shows the same match, but also including a measured brightness comparison. Blue and Green circles are potential/members as determined by the later stages of processing.

Since the stars categorized as non-member in the previous literature were already preselected as being in close spatial proximity to the members, I felt that using them as a negative labeled input would lead to flawed results (the negative sample would not be randomly distributed in both positional, proper motion, and parallax). Therefore, my training set consists of only positive labeled input. However, similar to Elkan & Noto (2008), we have a well-defined set of members, randomly selected from the unknown total population. Note that the random selection criteria is **not** met had I also considered luminosity/brightness/stellar type, due to selection bias. However, the existing members represent a random distribution in terms of spatial and kinematic criteria. I also expect that the initial training set will contain a population of mis-categorized stars, specifically "members", which when analyzed with the GAIA kinematics and parallax measurements, will be found to be non-members. As shown in Section 3.2 and Figure 3.3 nearly all of the incorrectly categorized stars are eliminated during the categorization process.

The existing member list is considered a "labeled" list of "positive" categorizations. For the unlabeled data, I select all available sources in the GAIA DR2 database, which have parallax and proper motion measurements, and a minimum of 75 good (astrometric_n_good_obs_al > 75) measurements. To limit observational errors in position, parallax, and proper motion, I select only stars with GAIA G magnitudes brighter than 19.0 ($G < 19.0$). For clusters in extremely crowded regions, or clusters which cover a large angular region (like the Hyades), the number of sources returned by the GAIA database exceeds the data cap (~3M sources), so I add in an additional restriction on the parallax. Table 3.1 lists my ten target clusters, the membership reference used to form the training set, the angular search radius used for the GAIA DR2 data set, and any additional restrictions I used for the search parameters.

TABLE 3.1: Model Data Set Selection

Cluster ID	Prior membership source	Search radius	Data set size	Additional restrictions
Coma	Kraus & Hillenbrand (2007)	15°	32 (MB)	$Parallax > 3.0$
M7	Simbad	5°	552 (MB)	
Praesepe	Kraus & Hillenbrand (2007)	5°	106 (MB)	
Hyades	Perryman *et al.* (1998)	20°	169 (MB)	$Parallax > 2.5$
IC 4756	Missana & Missana (1995)	5°	629 (MB)	$Parallax > 0.8$
NGC 752	Platais (1991)	5°	191 (MB)	
NGC 3680	Anthony-Twarog *et al.* (1991)	5°	202 (MB)	
IC 4651	Meibom, Andersen & Nordström (2002)	1.5°	252 (MB)	$0.2 < Parallax < 3.0$
M67	Geller, Latham & Mathieu (2015)	5°	108 (MB)	
NGC 6791	Tofflemire *et al.* (2014)	3°	500 (MB)	

3.1.2 Gaussian Kernel Density Estimation

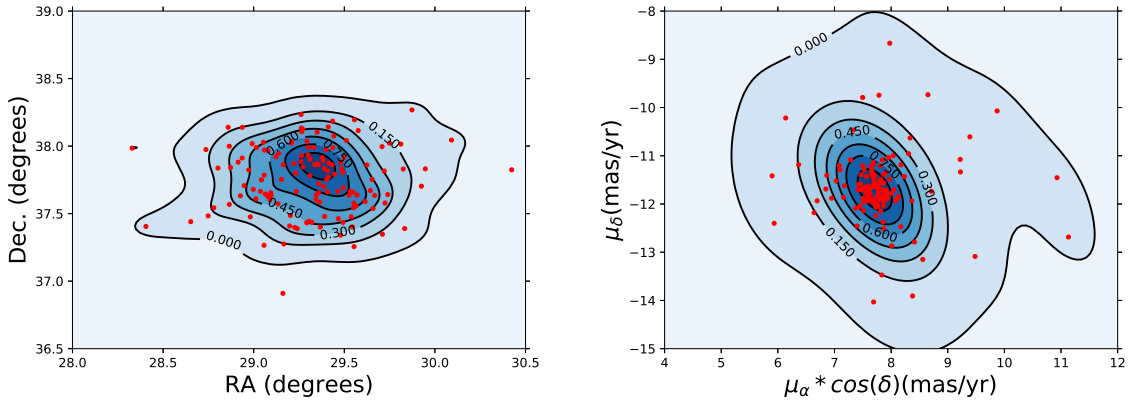
My membership classifier uses a weighted (average) combination of the probability distributions for three[2] parameters for each data point. For each parameter, I create a Kernel Density Estimation (KDE) function using the parameter uncertainties as the broadening parameter. The classifier then calculates a membership score for each "unlabeled" point in the GAIA DR2 dataset by averaging the KDE values for each parameter. For numerical consistency, I normalized each KDE to a peak value of 1.0, which allows for a simple weighting and scoring algorithm. The weights for the parallax, and proper motion parameters are set to 5.0. The weight for the positional component varies with the mean parallax, in order to allow for nearer clusters to subtend a larger area of the sky. For nearby clusters (parallax > 10), the positional component is weighted at one tenth the weight of the parallax and proper motion components. As the cluster distance increases, the weight of the position component increases until it reaches a value of 8.0 for clusters with parallaxes below 0.5.

Figure 3.2 illustrates three KDEs used for the membership scoring calculation of an example cluster (NGC 752), as well as a fourth KDE for the radial velocity parameter. The data used for all KDEs is the corresponding data from the GAIA DR2 dataset for the source which most closely matches the RA and DEC of the literature member. For reference, I have overlaid a Gaussian distribution (in red) over the KDE function (black) in the two 1-D plots. The two 1-D estimators are formed by summing Gaussian distributions, centered on the relevant measurement (either parallax or radial velocity) for each star, and using $\sigma = \epsilon$ of the relevant measurement's error.

While I have implemented the functionality to include the radial velocity parameter as part of the membership classifier, the current state of the radial velocity parameter in the GAIA data set precludes that parameter from current use. The radial velocity parameter is not available for sources dimmer than $G_{RVS} = 14.0$, or outside of T_{eff} range of 3550–6900 K. The result of the limited sample of radial velocities can be seen in Figure 3.2d where the best-fit Gaussian function is significantly different from the (insufficiently) broadened KDE.

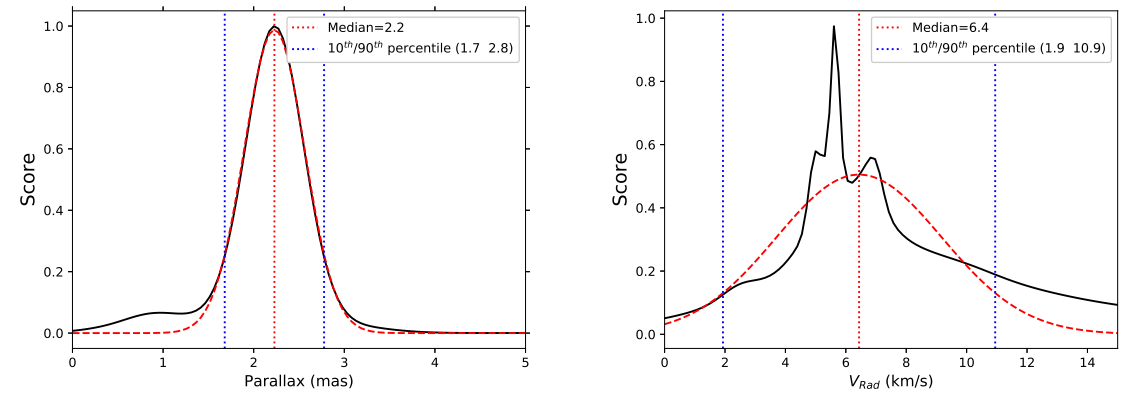
[2]Technically, a probability distribution for the radial velocity parameter is calculated, and included as a fourth probability distribution. However, it is assigned a weight of zero (0.0). This is in anticipation of the quality of the radial velocity parameter improving in future releases.

The two 2-D estimators for proper motion and position (Figures 3.2a,3.2b) are calculated in a similar manner. However, due to the precision of the Gaia measurements for these factors, if only the combination of measurement and equipment uncertainty is used as a bandwidth/broadening function, the resulting KDE is under smoothed. Therefore, I opted for a more generalized broadening function using "Scott's" rule (Scott 1992), as implemented by the *SciPy stats* package.



(A) Position KDE

(B) Proper motion KDE

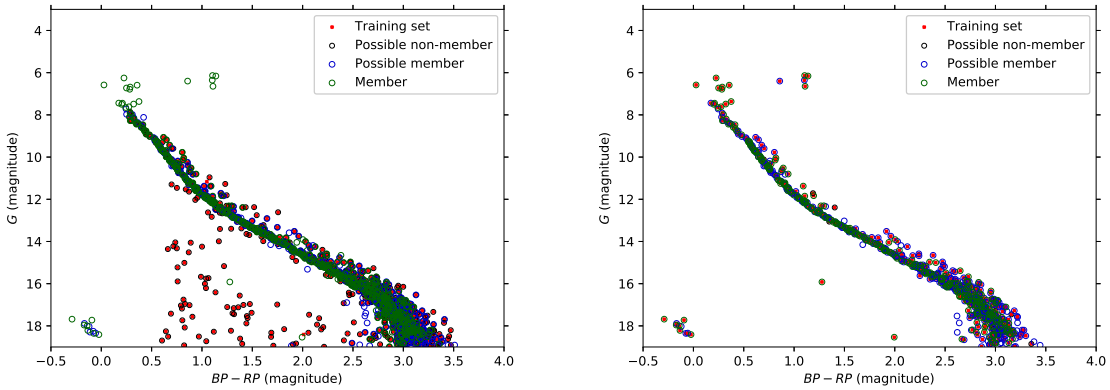


(C) Parallax KDE

(D) Radial velocity KDE

FIGURE 3.2: The four Kernel Density Estimators (KDE) used to establish Bayesian priors for calculating the membership score. Data shown is for NGC 752. Note the non-Gaussian form for the radial velocity parameter (3.2d). The Gaussian fit (red) is broadened, and will cause stars which are not cluster members to have inflated scores from this KDE.

3.2 Color-Magnitude Diagrams and Process Evaluation



(A) Praesepe CMD after one iteration

(B) Praesepe CMD after two iterations

FIGURE 3.3: Side-by-side comparison of the Color-Magnitude Diagram of Praesepe. Figure 3.3a (left) contains the initial (Kraus & Hillenbrand 2007) member list in red, the stars rated as cluster members in green, and probable members in blue. Stars, listed as members in the literature, but with scores which do not reach the probable member threshold are circled in black. The training set in Figure 3.3b (right)are the high-likelihood members from the first iteration.

In order to determine when to terminate the iterative classification process, I chose to use a color-magnitude relationship. Since the observed GAIA magnitude values are not used in creating the generative model, they provide an independent indicator of the success of the classification process. Conveniently, the the color-magnitude relationship for clusters is relatively well understood, and can be discerned by eye as a clear main sequence. Figure 3.3 shows the evolution of the CMD for the Praesepe cluster. When mapping the Kraus & Hillenbrand (2007) members to their GAIA DR2 counterparts, there are a large number of cluster members at $G > 17.0$ which map to points to the left (blue-ward) of the apparent main sequence. Their position on the CMD is strong evidence that they have either been mis-categorized as cluster members, or have been mis-matched to GAIA sources. The fact that the parallax and proper motion components of these members correspond to their fellow Praesepe members, indicates that the mis-categorization is the more likely explanation. This CMD has the literature training set overlaid in red, potential cluster members from the unlabeled GAIA data in blue, and stars classified as members in green. Downgraded members, or literature members which the classifier has scored below the minimum for potential membership, are red points, circled in black.

As a result of having the questionable, or mis-identified, members in the training set, the first iteration of the classifier produces a number of similarly (mis-) placed points. However, most of these points have low membership scores, and are not included in the training set for the second iteration. The subsequent CMD Figure 3.3b is the CMD after the second iteration. Note depletion of the likely non-members as a natural consequence of the classification process. For four clusters which are either more distant (NGC 3680, NGC 6791), or lie in

TABLE 3.2: Cluster Member Statistics

Cluster ID	Training set size	GAIA DR2 set size	Deprecated members	# Members (GAIA) Probable	Possible
Coma	150	14673	83	101	49
M7	221	818606	104	132	402
Praesepe	1045	420598	417	354	494
Hyades	218	626501	112	74	102
IC 4756	99	2400007	5	63	228
NGC 752	135	349674	10	53	149
NGC 3680	50	778951	6	20	37
IC 4651	95	944823	5	32	221
NGC 2682 (M67)	695	431873	117	119	655
NGC 6791	170	162409	27	378	681

a crowded region of the sky (M7, IC 4651), a third iterative step was needed to produce a cleaner main sequence.

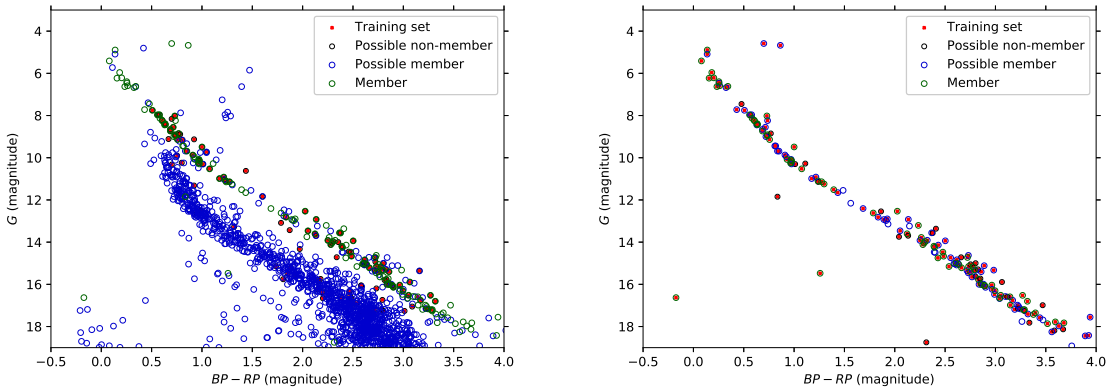
I include the full set of KDE (before/after) image pairs, and the final CMD for each of the eleven clusters in Appendix B. A quick summary of the training set size, and final count of members and possible members is in Table 3.2. Finally, the listing of individual stars, coordinates, membership scores, and other statistics are available in digital format.

3.3 Cluster Membership

Final membership scores are listed by cluster and star in Table A.1 (Included as a digital table). These tables represent the stars with highest membership scores, and it is important to restress that the scores do not represent a membership probability.

For the ten clusters in common with our spectroscopic study, I only use our membership evaluation as a verification process. I double checked that any spectra I selected belongs to a star in the target cluster. As I stated earlier, the choice by the original observers to measure particular stars in the region of our target clusters naturally pre-supposes cluster membership. However, by using this process, I will be able to easily reference a much larger set of membership data for any future cluster studies.

3.3.1 The Curious Case of Coma



(A) Coma CMD after one iteration

(B) Coma CMD after two iterations

FIGURE 3.4: Coma CMD after the first (left), and second (right) classifier iterations. After the first iteration, the classifier appears to have found two populations, one being the expected population of stars in the Coma cluster, and a second lying on was appears to be a main sequence. The second population is removed after the second iteration of the classifier, due to the stronger correlation of the original training set with the known Coma cluster characteristics.

Figure 3.4 shows the CMD for the Coma cluster, based on categorization from the membership list in Kraus & Hillenbrand (2007). The majority of the training set (red), and the GAIA sources identified as "members" lie on a main sequence. As with Praesepe in Figure 3.3, M7 (Figure B.9), IC4680 (Figure B.5), and others, there is a subset of training set members to the right of the main sequence. It's unlikely that this population are cluster members - probably being mis-categorized as possible members due to larger errors in the GAIA DR2 data with the dimmer and more distant stars. In the Coma data, however, these stars do not appear to have a random distribution, rather they appear to form a well-defined main sequence. Furthermore, Coma is relatively nearby (86pc) where, similar to our Hyades analysis, the non-member stars do not pollute the sample population. Examination of the parallax model shows a secondary concentration of member candidate stars at a parallax of approximately 3.5, in addition to the expected concentration of stars at the known parallax for the Coma cluster of 11. In fact, as shown by the progression of the parallax KDE (Figure B.2), the generative algorithm briefly believes that the parallax 3.5 population is the desired result.

While the generative process eventually settles on the "right" result for the known Coma cluster stars, the other population looks enough like a second, background cluster to merit further investigation. While I briefly toyed with the idea of the discovery of a previously undiscovered cluster, there were a number of pieces of evidence to the contrary. First, when observing the spatial and proper motion distribution of these "cluster" stars, they were far more diverse than those of the other clusters. Figure 3.5 shows what the 2-d KDEs look like for a sample of the mystery "cluster" stars. The cluster has an approximate diameter of 10 degrees. The apparent correlation of background stars is simply the "natural" main-sequence of Milky-Way stars. Figure 6 (reproduced in Figure 3.6) in Gaia Collaboration *et al.* (2018)

illustrates three color-magnitude diagrams, taken from stars in the immediate Solar neighborhood. When corrected for the distance modulus of 7.3 (parallax=3.5), my apparent "cluster" matches exactly with the random sample of Solar neighbors.

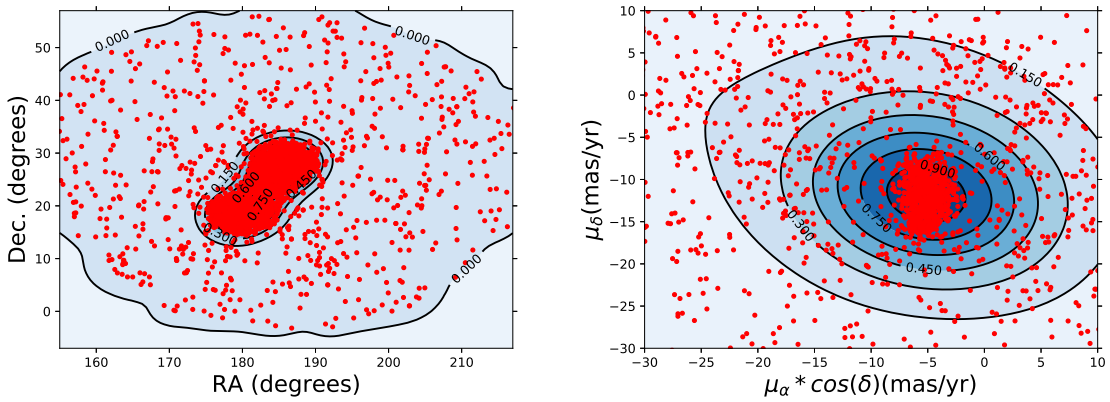


FIGURE 3.5: The position and proper motion distributions of the background "cluster" in the Coma field. These distributions are too widely dispersed to be a gravitationally-bound open star cluster.

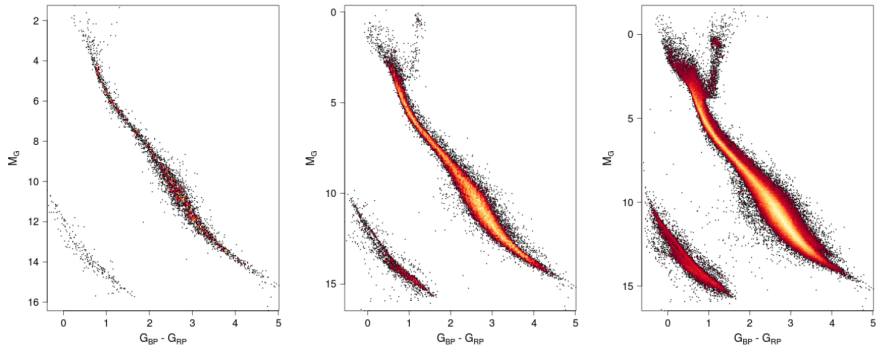


Fig. 6. Solar neighbourhood *Gaia* HRDs for a) $\varpi > 40$ mas (25 pc, 3,724 stars), b) $\varpi > 20$ mas (50 pc, 29,683 stars), and c) $\varpi > 10$ mas (100 pc, 212,728 stars).

FIGURE 3.6: Figure 6 from Gaia Collaboration *et al.* (2018), plotting the population of Solar neighbors.

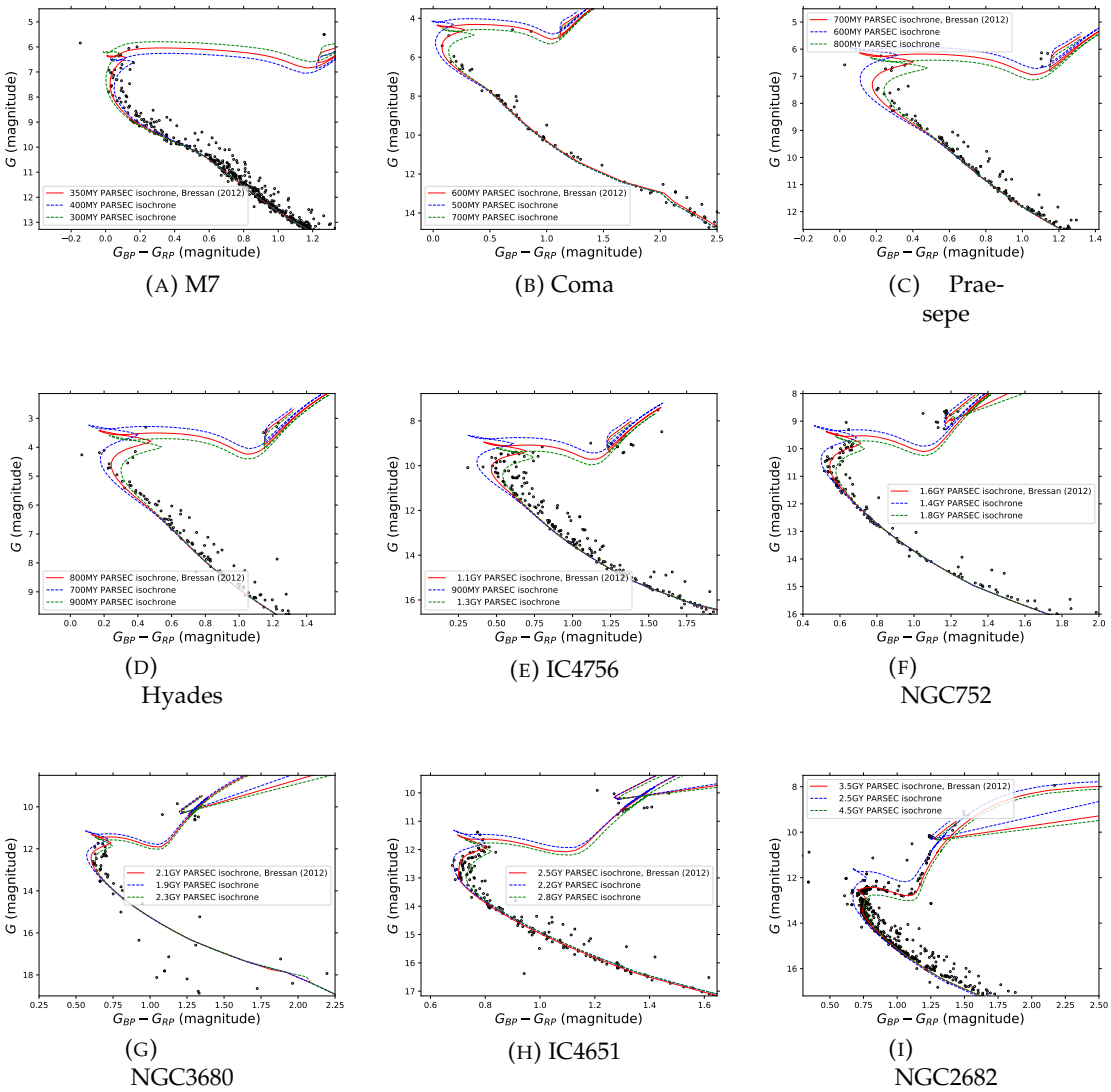
3.3.2 Algorithm Enhancements and Continuing Development

It seems that no software project is ever complete, and this analysis package is no exception. I foresee future releases incorporating a replacement for the manual CMD main-sequence stopping condition. As a result of the classification process, I obtain a reasonably accurate parallax and, therefore, cluster distance. By using a properly fit isochrone, I can develop a statistical evaluation (possibly a χ^2 evaluation). An automated stopping condition check removes the only subjective measure in the process.

Once more accurate radial velocity data is available, I can (re-) activate the fourth KDE parameter, including the full kinematic information for any star in the classifier.

Finally, there are hundreds of clusters which I did not evaluate. While I had fully expected the GAIA team to have published a cluster membership study, similar to this one, as of this writing (May 2018), no results have been forthcoming. As the membership results from this process are mainly due to the quality of the GAIA satellite data, I would defer any complete cluster membership catalog publication from GAIA data to the GAIA science team. However, barring any such publication in the immediate future, I would like to undertake such an effort.

3.3.3 Cluster Parameters from Isochrone Fitting



(A) M7 (B) Coma (C) Praesepe

(D) Hyades (E) IC4756 (F) NGC752

(G) NGC3680 (H) IC4651 (I) NGC2682

FIGURE 3.7: Isochrone fits for nine clusters

I used the on-line CMD isochrone generator tool (v3.1) `http://stev.oapd.inaf.it/cmd` (Marigo *et al.* 2017) to produce isochrones, which I then combined with GAIA photometric data for

my cluster members to produce the age, distance (through the distance modulus), and interstellar reddening parameter. Figure 3.7 shows the Main-Sequence Turn Off (MSTO) region of nine clusters, including seven from this work. The MSTO point for NGC 6791 lies too close to GAIA DR2's limit of G=18.0 to be well fit with an isochrone. Table 3.3 summarizes my cluster parameters from isochrone fitting, below.

TABLE 3.3: Cluster Member Statistics

Cluster Name	Age (MYr)	Dist. Mod.	Dist. (pc)	E(G_{BP}-G_{RP}) (mag.)
M7	350	7.3	290	0.08
Coma	600	4.7	87	0.00
Praesepe	700	6.3	180	0.00
Hyades	800	3.4	48	0.00
IC 4756	1100	8.4	480	0.12
NGC 752	1600	8.3	460	0.04
NGC 3680	2100	9.3	730	0.05
IC 4651	2500	9.7	870	0.08
NGC 2682 (M67)	3500	9.8	910	0.05

Chapter 4

Software

The process of calculating stellar elemental abundances is a complex one. The astronomer must obtain raw high-resolution spectra, correct the spectra for instrumental and wavelength (Doppler) effects, flux normalize the spectra, measure absorption features in the spectra, and then re-create (model) the atmospheric conditions (composition, temperature, density/gravity, turbulence) which would best represent their measurement. While observers no longer have to perform all of these steps by hand (including tracing the spectra!), the software packages available to assist in spectra processing and abundance calculation is still relatively immature.

When I embarked on this project, I knew that I would have to provide software automation to "fill in the gaps" where current software either produced an "intermediate" product (such as a reduced spectrum), or where no automation was available at all. The remainder of this chapter is a documentation of the various software functions I have implemented, with digressions at the appropriate times to discuss design considerations, statistical methods and database organization.

In order to provide open-source code, and also in the interest of creating a fully reproducible process, I maintain all of my code in a public repository at:

`https://github.com/mikelum/ClusterAnalysis`

The code is organized as a series of inter-dependent modules, loosely based on the stellar abundance process as detailed in the following sections. As much as I would like to have this code widely adopted for use in the field, I expect that it will not be widely used beyond this work. However, in the off chance that it will be used by others in the future, I have followed the standards in the "Style Guide for Python Code" (van Rossum, Warsaw & Coghlan 2013), and provided extensive in-line comments, function descriptions, and provided references where applicable.

4.1 Overall Design Decisions

Generally speaking, there is no standard for language, database software, operating system, image formats, or pretty much any other digital processing/analysis of astronomical data. While there has been a convergence toward using the .fits (Wells, Greisen & Harten 1981) format for astronomical images, I had to make decisions early on in my research process as to which standards to use. The most important of these would be the programming/scripting language, and the database format.

A non-language consideration I feel I should note is that, when presented the choice between code readability and efficiency, I always opted for the more readable path. While more efficient code is more desirable for calculation intense functions which run often, my code is not intended to run continuously. Rather, I expect that there will be a relatively small number

(<1000) of stellar spectra processed for a given project, and that the total amount of processing time would be on the order of a day or two on a typical desktop computer. I estimate that there are a significant number of optimizations possible in the existing code, enough to potentially reduce processing of a typical spectrum by a factor of three to five. However, optimization would come at the cost of readability and maintainability of the code. I believe that the potential future appeal of the code will not be in its efficiency, but rather in its accessibility, thus it would be more important to reduce the learning curve for future coders.

4.1.1 A note on the use of Python

With my background in software engineering, I have extensive experience in programming languages and could choose my software language, based on the "best" considerations, rather than programmer experience. For this type of application, I felt the main options to be Fortran or C, as low-level, high-efficiency options, vs. IRAF[1] and Python[2] as less-efficient options, but with far wider use in the astronomical community with the accompanying wider selection of existing libraries.

I opted to use Python partially due to its growing popularity — it was rising in the list of top languages in 2010, and has now topped IEEE's list of most popular languages for open source software development for two years running (Cass 2018). Python also has extensive, free/open source numerical analysis, scientific support, and astronomy-specific libraries (e.g.: NumPy, SciPy, and AstroPy, respectively, to draw out the major players). The IRAF language is the only other option with a significant number of supported and relevant libraries, but the number of available IRAF libraries is an order of magnitude fewer than those in Python. Additionally, IRAF was[3] a commercial product, and was not freely available.

Finally, and as the factor I consider most important, Python is an interpreted language; designed to be easily readable. It is, therefore, portable to any operating system which has a Python interpreter (which includes all major operating systems, MacOS, Windows, and nearly all "flavors" of Linux and Unix). Readability directly correlates to maintainability, meaning that should others wish to use my software, their learning curve will be greatly reduced.

4.1.2 SQLite Database

For many of the same reasons, I opted to use the SQLite[4] version of the Structured Query Language "SQL" as my database format. I made this decision primarily due to the fact that SQLite is a freely available, open-source product, with a corresponding freely available Python (interface) library. As an added bonus, SQLite is a Recommended Format by the Library of Congress for long-term preservation and access[5]. I will go into greater detail of my database configuration, below.

[1] IRAF is distributed by The Association of Universities for Research in Astronomy, Inc., under cooperative agreement with the National Science Foundation

[2] "Python" and the Python logos are trademarks or registered trademarks of the Python Software Foundation, and are used herein without prior approval, under "nominative use rules."

[3] In 2012 NOAO removed all license restrictions, making IRAF "free"

[4] SQLite is public domain software supported by the SQLite Consortium `https://www.sqlite.org/index.html`

[5] `https://www.loc.gov/preservation/resources/rfs/data.html`

4.1.3 Scripting and External Interfaces

The interface method to much of the existing astronomical software, like the MOOG (Sneden 1973) atmospheric modeling package, and IRAF-based image processing and display functions, is a relatively primitive command-line-interface. Conveniently, Python provides a command-line scripting interface. So, much of the interface between these external functions and my database and processing functions is seamlessly performed by Python routines. Because of the requirements of the external functions, I cannot claim my software is fully cross-platform compatible, but for my sample size of three Linux machines, running both versions 16.04 and 18.04 of the Ubuntu[6] "distribution" of Linux, it is fully functional.

4.2 Process Walk-Through

While most steps in my software product are stand-alone, or re-entrant, the best way to describe its operation is to provide a step-by-step walk through of the process of taking a spectrum from archive through determining its star's elemental abundance measurements.

The outline of the process looks like:

1. Spectra Acquisition (Section 4.3)
2. Spectra Processing (Section 4.4)
3. Absorption Line Measurement (Section 4.5)
4. Stellar Parameter Determination (Section 4.7)
5. Abundance Calculation (Section 4.8)
6. Output/Assessment — Tables and Plots (Section 4.9)

4.3 Spectra Acquisition

In a perfect world, obtaining spectra from archives would be a simple process. The user would enter the name or coordinates of the object they wish to observe, and the archive would return a properly wavelength- and flux-calibrated spectrum with the highest resolution and signal-to-noise available. Sadly, the reality is different. While most archives are moving to releasing spectra which are closer to the end product needed for science analysis, all of the spectra used in this project required some sort of intermediate processing.

4.3.1 Manual Archive Searches

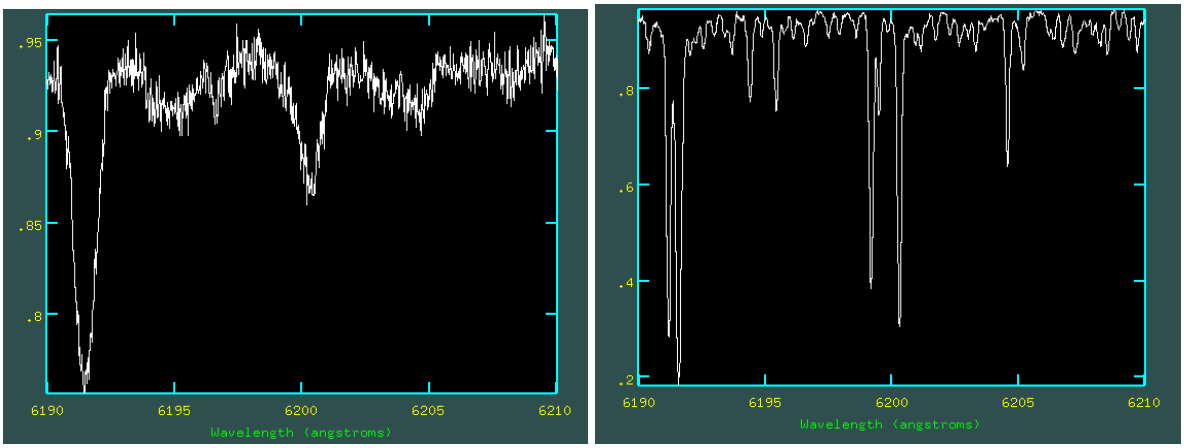
I obtained spectra from three main sources, all three of which have online interfaces. While I manually selected the 800 or so spectra used in this study from the available data, Python libraries, like the urllib package, would allow the request and response to be automated for larger projects. For an example of how online form queries and response parsing would work, see the next subsection where I detail photometry and cross-reference lookups into the Simbad (Wenger *et al.* 2000) database.

[6]Ubuntu is a registered trademark of Canonical Ltd. Distribution and use of Ubuntu software is free, and governed by the Linux Open-Source Agreement

The three archives used were: The Keck Observatory Archive (KOA) `https://koa.ipac.caltech.edu/cgi-bin/KOA/nph-KOAlogin`, The SOPHIE & ELODIE archive at the Observatory of Haute-Provence `http://www.obs-hp.fr/archives.shtml`, and the European Southern Observatory archive `http://archive.eso.org/wdb/wdb/adp/phase3_main/form`. Note that the ESO archive form linked here is now obsoleted, in favor of the programmatic, and GUI versions at: `http://archive.eso.org/programmatic/#INIT` and `http://archive.eso.org/scienceportal/home`, respectively. The spectrographs which met the R>40,000 requirement were HIRES (Vogt *et al.* 1994) for the Keck Observatory, HARPS (Mayor *et al.* 2003), FEROS (Kaufer *et al.* 1999), and UVES (Dekker *et al.* 2000) from the ESO archive, and SOPHIE (Bouchy & Sophie Team 2006) and ELODIE (Prugniel & Soubiran 2001; Moultaka *et al.* 2004) from OHP.

I used a consistent selection process for all three archives. For each cluster, I requested a listing for all optical spectra for objects in a radius around the central coordinates of the cluster. For clusters which subtend a larger angle on the sky, like the Hyades and Pleiades, I used a larger search radius (5°). For the more distant clusters, I used a smaller, 1° radius. When given a choice of instrument configurations, I selected for only high-resolution spectra (R>40,000), with no radial velocity (iodine) filtering. I then selected only stellar spectra which matched the "optical" wavelength criteria (a significant portion of the spectra was in the 450-850nm range), and which matched coordinates to stars with a membership "score" of 50 or better for the given cluster (see Chapter 3).

The final step in spectra selection was to perform a simple (manual) visual inspection. The inspection proved necessary to eliminate spectra which were mis-labeled as stellar spectra – extended object spectra which were labeled for the guide star — stellar spectra of white or brown dwarfs, or spectra like that shown in Figure 4.1a, of stars with high rotational velocities with excessively broadened absorption features.



(A) IC 4651, MEI-11799 (B) IC 4651, MEI-8540

FIGURE 4.1: IC 4651 stellar spectral samples from the ESO's UVES spectrograph at Cerro Paranal. The high rotational velocity spectrum from MEI-11799 (left) shows significant broadening (and the accompanying blending) of the absorption features which would be otherwise present in a low rotation velocity star, like MEI-8540 (right).

4.3.2 Automated Searches and Lookups

While spectra searches only required a couple of dozen searches, and could be done manually, I needed to create an automated lookup function to obtain both photometric data and to cross-reference the catalog indices I used to identify the stars in my survey, with the (up to) 20-digit ids used by the GAIA DR2 release. Since there is really only one centralized source, which contains both the standard set of Johnson-Cousins filters (UBVGRI), the 2MASS infrared set (JHK), and the GAIA DR2 filter (G) and "source" id, I opted to access the online data from the Simbad (Wenger *et al.* 2000) database.

Luckily, online inquiries into the Simbad database are encoded in plain text into the requested url. A sample url for HER-40 in IC 4756 looks like:

`simbad.u-strasbg.fr/simbad/sim-basic?Ident=Cl*+IC+4756+HER+40`

The code used for Internet lookups is located in the "InternetPhot" library within the photometric atmospheric parameter determination module ("ParamDet_Phot"). The Application Programming Interface (API) for these function accepts a cluster and star ID, and returns a "dictionary" of available photometric indices.Alternatively, the functions can return the Simbad Right Ascension (RA) and Declination (DEC) coordinates (ep= J2000), or the GAIA DR2 "Source ID" (if available). Internally, the code parses the webpage (HTML) code returned by the url request, looking for certain indicators for the desired data. Potentially, should the Simbad web page or data format change, these indicators may no longer be valid. However, the effort to code a more robust content-based parser would far exceed the effort reduced by the automation of the lookup.

4.4 Spectra Processing

In the "old-school" model of spectra observations, the astronomer is responsible for obtaining "calibration" images, and correcting the raw image from the telescope to remove any instrumental effects from their spectra. With the exception of Keck, over the past decade observatories have moved to using image reduction "pipelines" which provide "science-ready" images after automatically correcting the raw images, relieving the scientist of the engineering aspects of that process. In the case of all three of the archives used in this project, the observatories offer pre-reduced data. However, the Keck Observatory Archive offers the warning that: “the quality and content of the data...may not necessarily be suitable for publishable science.” (`https://koa.ipac.caltech.edu/UserGuide/HIRES/extracted.html`). While I acknowledge the KOA warning, I still opted to use the Keck data as provided. My reasoning being that, should I have chosen to do my own reduction, I would have used the same tools — specifically *MAKEE*, choice of calibration frames (as suggested by the archive search), and process as the automated pipeline. I do not have sufficient knowledge of the observation session to override the recommendations of the archive.

Once I had a set of stellar spectra, two more steps were required to obtain a "science" image. First, I applied wavelength and flux corrections and calibrations to each spectrum. At this point, I would also convert all spectra to a "1-D .fits" format, if they were not already in that configuration. Secondly, in cases where multiple images were required to obtain sufficient signal (or signal-to-noise ratios), I co-added (stacked) the series of images for the target. In cases where spectra for the same star were obtained from multiple instruments, I would maintain separate spectra from each instruments, and measure the (stacked) images from each instrument separately, taking the mean of the absorption line measurements.

For the wavelength (Doppler) correction, I wrote a Python function which calls the STSDAS[7] (IRAF) "splot" function and prompts a user to locate strong Fraunhofer and iron absorption features. It then calculates the overall velocity correction required for the spectrum, and applies the Doppler correction using STSDAS's (IRAF) "dopcor" function. Python to IRAF interactions were performed by the "PyRAF" package. I feel that it is noteworthy that the Doppler corrections for most of the KOA spectra implied a velocity correction of around 100km s^{-1}, which is significantly more than what would be expected from the Earth's orbital velocity ($\approx 30 kms^{-1}$) and the highest radial velocity of nearby clusters (generally <$50 kms^{-1}$).

The next step in my reduction process was to normalize the continuum flux of all spectra to a value of 1.0. I used a 5^{th} order polynomial fit for each spectral order, which I found was the highest-order polynomial I could use to fit the relatively smooth regions of the longer wavelength orders of the spectra, without over-smoothing the broad lines, and "noisier" regions. The final step applied to cases where multiple spectra from a single instrument were available for a single star. I co-added these spectra using a weighted mean, using STScI's IRAF function *scombine*, with each weighted by its signal-to-noise ratio.

4.5 Absorption Line Measurements

Using the continuum-fit and Doppler-corrected spectra, my software then attempts to measure approximately 2700 absorption line features per spectrum from a list of 31 elements, 8 of which have two ionization states. A complete listing of the line list (2659 lines), reference list (68) and equivalent width measurements (382288) for all stars are available as supplemental digital data. Appendix C contains the entire line reference key table (in order to have the sources included in the bibliography), and samples and formats from the longer tables are in Appendix A.

The automated line measurement algorithm is a combination of Python scripts and *AstroPy* functions. To obtain EQuivalent Width (EQW) measurements, the software first builds a custom line list for each spectrum, based on its wavelength coverage. The software establishes a baseline continuum of the (0.4nm wide) spectral region around a given line to account for local line blanketing (continuum reduction by molecular and metallicity effects), and for systematic offset by the general continuum fit from the reduction process. The measured equivalent width is set to the area between this (local) baseline continuum and the best curve fit to the spectra points in that region. The curve used is a simple linear combination of Lorentzian and Gaussian curves. The "best fit" line center is set to the local minimum nearest the expected line center, with the range of the fit set to the two local maxima which bracket that local minimum. For blended lines, either (or both) of the local maxima can lie significantly below the local baseline (continuum. In these cases, the Gaussian/Lorentzian curve fit will extend to cover points outside of the initial range (See Figure 4.2).

While our fitted curve consists of the Gaussian and Lorentzian linear combination, we found that, after trying multiple line profiles (pure Gaussian, Voigt, and rudimentary linear and trapezoidal interpolations), our chosen line profile had less effect on the calculation of the equivalent width of a given line than individual data point errors from signal-to-noise, and spectral dispersion. However, we opted for the more complex fitting curve out of respect to the physics of line generation, specifically the Gaussian core and Lorentzian wings.

[7]STSDAS and PyRAF are products of the Space Telescope Science Institute, which is operated by AURA for NASA

For our NGC 752 publication, we compared widths measured by our automated algorithm to those taken by hand in other publications. Figure 5.2 compares over 1200 IRAF *splot* equivalent width measurements taken by Reddy, Giridhar & Lambert (2012) and Böcek Topcu *et al.* (2015). We found no systematic difference between the techniques ($\Delta_{EQW} = 1.2 \pm 5.9$ mÅ), nor was there a significant difference between the variance between our method and either of the prior works, or between the two prior works. Specific outliers are discussed further in Chapter 5.

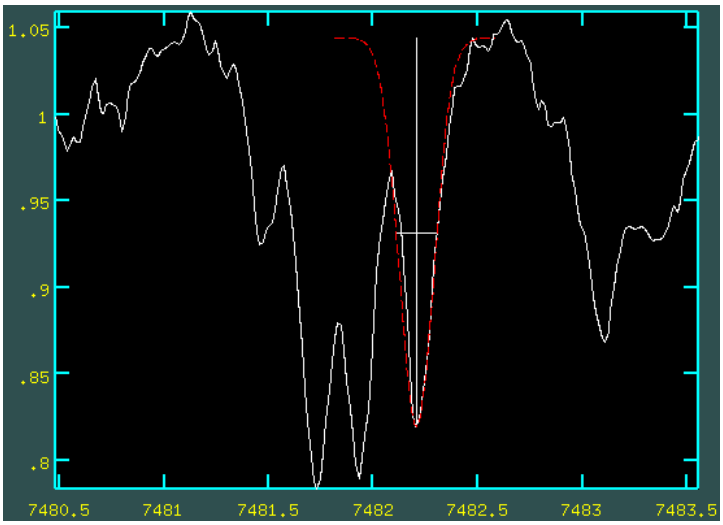


FIGURE 4.2: Fitting a blended line. The Gaussian core of the line is fit to the local minimum at 7482.2Å, with the range of the fit extending to the local continuum (near 1.05). The spectral points fit range from the local maximum at 7482.1Åto the one at 7482.4Å. Since the fitted curve extends beyond the range bounded by the two local maxima, the equivalent width measurement will include portions of the nearby "blended" lines.

4.5.1 Equivalent Width Filtering

The (SciPy) curve fitting algorithm used to fit absorption features within a given spectrum obviously has no knowledge of what to expect as a "typical" feature line. This issue manifests in two categories of erroneous fits. When a line measurement falls into one of these categories, my software must first detect the "bad" fit, and then either opt to re-fit the line or, failing to do so, exclude that measurement from being entered into the database.

The first category is when lines are too "small." For many lines in my line list, the resulting absorption can be so small, that it is indistinguishable from noise "features" along the continuum. As there is no default minimum fit size, the software will occasionally attempt to fit features which are "not real". To avoid these cases, I implemented two filters. The first is in the initial fitting algorithm. If the two bracketing (local) maxima are too close (<5 pixels, or 0.05-0.10Å, depending on dispersion), the algorithm widens its fit region to the next two local maxima. The second filter is for the small lines which make it past the 5-pixel minimum width. As there is no way to distinguish a weak signal from noise, I establish a minimum EQuivalent Width measurement (EQW) for any given spectra of ten (10) times the product of the resolution (in mÅ/pixel) and inverse of the signal-to-noise, with a minimum of 2.0 mÅ. By no coincidence, this value is the same as the minimum 5 pixel maxima-to-maxima width multiplied by twice the size of a "typical" noise feature as measured in the spectra with the highest dispersion (D=20mÅ).

The second category of mis-fits are (appropriately) when the line fit is too "broad". A first-order selection occurs during the abundance calculation process where I only accept absorption features which would lie on the linear portion of the curve-of-growth (detailed in Section 4.7.2, below). However, lines which lie on the non-linear portions of the curve of

growth are legitimate measurements, just not desirable for abundance calculations. In some cases, a combination of my software selection of spectral region and the χ^2 optimization of the curve fit, incorrectly fits a section of the spectrum. As shown in Figure 4.2, the fit is allowed to include points outside of the bounds of the two local maximum. For very "shallow" features, near a smooth continuum level, the fitting algorithm may decide that the best χ^2 fit subtends an extremely wide (>2Å) region. To eliminate these wide fits, I measure the line profile in terms of the ratio of the Full Width at Half Maximum (FWHM) of the Gaussian portion of the line core (see Section 4.5, above) to the resultant EQW measurement ($FWHM/EQW$). When our automated measurement process finds a potential absorption line, which was incorrectly fit to a broad section of random noise fluctuation in the spectrum, the $FWHM/EQW$ ratio is almost always below 1.0 ($>$ 99.8% for the lowest S/N spectra). I determined this ratio limit of 1.0 through Monte Carlo (MC) measurement simulations on 250 random noise spectra. Each noise spectrum point consists of a (normalized) flux value, generated around a continuum value of 1.0, with a variance determined by the S/N being modeled. I then inserted randomly selected regions from actual data, containing a single absorption feature, into the random noise spectra. I ran each noise spectrum through the line measurement process, and evaluated the resulting line measures. I applied the same criteria for line acceptability (within the linear portion of the curve of growth, within the expected wavelength range, and above the detection threshold) to these line candidates, eliminating approximately 80% of the (false) lines from a noise spectrum prior to applying the line profile criterion. I was able to recover 100% of “real” lines inserted into the test spectra, using the 1.0 (FWHM/EQW) ratio limit, and incorrectly classify less than 0.2% of the “false” lines as legitimate measurements. Of the incorrectly measured lines, an average of 3 per Monte Carlo run of 2700 lines in 250 simulated spectra passed both the 1.0 ratio limit and the line acceptability criteria, as listed above.

4.6 Database Formats

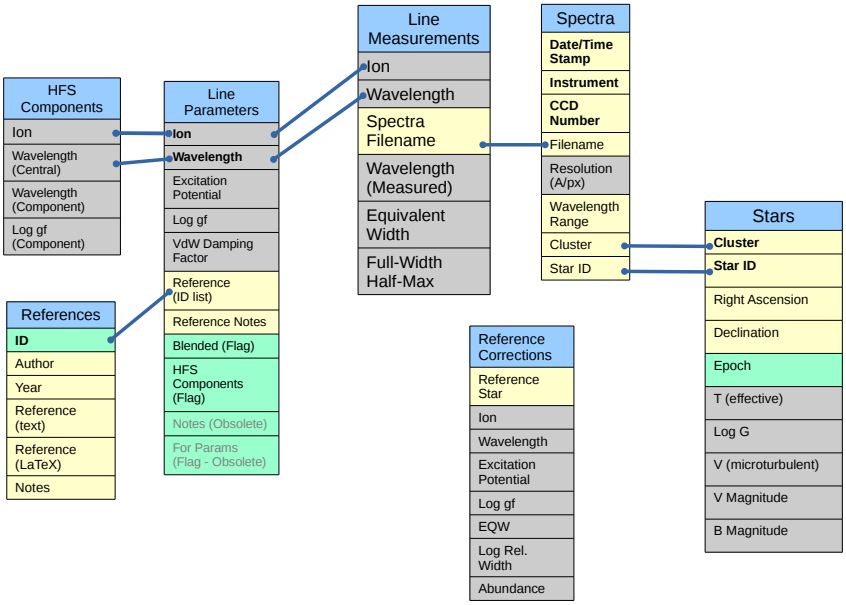


FIGURE 4.3: Schema of the database used for this project. Table/field cross references are labeled with blue lines. Primary keys for each table are in **bold**, with background colors representing the data type — yellow=text, green=integer, gray=real

Even though the majority of the data products used for this project (i.e.: spectra, photometry, and absorption line physical parameters) are available from external sources, the most efficient way to utilize this data is to maintain local copies. In addition, intermediate products, like the previously discussed EQW measurements need to be stored, to avoid repeated measurements every time a subsequent calculation requires that data. Figure 4.3 lays out the database format ("schema"), with the inter-table "relations" (thus "relational" database) noted by dashed lines. Each "Table" is addressed individually in the subsections which follow. Unique keys (identifiers) within each table are highlighted in **bold**, while actual data types are noted by background shading (yellow=text, blue=integer, gray=real). In the text discussion below, names of data fields are *italicized*.

4.6.1 Spectra File Information

After the reduction process is complete (Section 4.4), as a prelude to the EQW measurement process, my software will enter information for each spectrum into the database. Each spectrum is uniquely defined by the *date and time* it was recorded, and the *instrument* it was recorded on. The *CCD* (number) used was previously recorded, but that field is currently unused and only remains for archival reasons. Data for a given spectra are the actual *filename*, the *cluster* and *star ID* of the target star, and the *resolution* and *wavelength range* of the spectrum.

4.6.2 Stellar Parameters

The *cluster* and *star ID*s also uniquely identify the star in a second table. While the primary keys for individual stars are the *cluster* and *star ID*s, each entry also contains *Right Ascension* (RA) and *Declination* (DEC) coordinates for the star at a given *epoch*. Each entry in the stellar parameters table also contains physical and observational parameters for *effective temperature*, *surface gravity*, and *micro-turbulent velocity* needed to create atmospheric models, and subsequently calculate elemental abundances. Photometric parameters for *V magnitude*, and *B magnitude*, as recorded in the Simbad database. The photometric data fields are only filled if that data is requested as a part of another process — it is not automatically populated when the stellar data is first entered.

4.6.3 Absorption Line Parameters and References

The process of EQW measurement detailed in Section 4.5, begins with a *wavelength* lookup into the database table containing physical line parameters to find all potential absorption features (lines) in a given spectrum. Each line entry is uniquely identified by its *wavelength* and *ion*. The contents for each entry consist of the physical parameters of that feature — its *excitation potential*, probability of transition ($Log g_f$), and a *vanDerWaals damping* factor. The list of *reference* source IDs for the parameters used is also included, and cross-references into the references table (Appendix C). Three additional fields are included to assist with abundance calculations. The *ForParams*, and *Blended* flags are used to select lines for atmospheric parameter determination, or to de-select it as a less-accurate "blended" line, respectively. Certain lines which have hyperfine structure are designated by a "1" in the *HFS* field. These lines will also have corresponding entries in the hyperfine table.

4.6.4 Hyperfine Line Data

Specific absorption features with known hyperfine structure will have entries in this table. Each feature has the *wavelength* and *ion* of the relevant line, as well as the *feature wavelength* and (fractional) $Log g_f$ of the feature. When calculating elemental abundances, hyperfine features are considered multi-element "blends" by the analysis and synthesis software.

4.6.5 Line Reference Data

"Credit where credit is due." The sources for the physical parameters for the absorption lines are stored here. Each is uniquely identified by an *ID* number, sequentially assigned as new lines are entered into the database. Each entry contains the *first author* and *year* of publication, with a full *BibTeX entry*, which can be directly copied and placed into a LaTeXbibliography. I felt the need to include generic *notes* field to annotate any special considerations used when including the data from a given source.

4.6.6 Measured Line Data

The largest table contains all of the measured EQWs for all of the measured spectra. Each entry contains the *wavelength* and *ion* of an attempted line measurement. Measured values for the central *(measured) wavelength*, *Equivalent Width (EQW)*, and *Full-Width at Half Maximum (FWHM)* for every detected line in a given spectra *file name*, are recorded here. Should an

attempted measurement not detect a line, a value of -1 is entered for *measured WL*, *EQW*, and *FWHM*. The absorption line data I selected comes from >60 sources. The full versions of Tables A.1 (Digital inclusion) and C.1 (Appendix C) have the physical parameters, and sources, respectively

4.6.7 Reference Star Measurements

The final two tables contain information used to calibrate the quality and accuracy of measured lines. For calibration purposes, I selected three stars as "reference" targets; the Sun, Aldebaran, and Arcturus. The first table contains the results of equivalent width measurements made for the calibration spectra. Specifically, it contains a list of *EQWs* measured for each line in the calibration *star*, as designated by *wavelength* and *ion*. I also maintain a line *delta* field as an indicator of the accuracy of a given line measurement across spectra.

The *abundances* calculated for the *reference stars* from each individually measured *EQM* (lines designated by *wavelength* and *ion*) are stored in the second "reference corrections" table. This table is mainly maintained as an alternative to calculating reference star abundances every time a reference comparison is required (See Section 4.8).

4.7 Stellar Parameter Determination

In order to translate the measured EQW values into a representative abundance value for an element in a given star, we used an external stellar atmospheric model simulation program (MOOG, see further discussion below), which requires as inputs, the conditions as represented by an effective temperature (T_{eff}), the ($\log_{10}$ of the) surface gravity of the star (LogG), and a "micro-turbulent velocity" (V_{turb}) parameter to represent line broadening.

I derived these conditions through an iterative process, based on spectroscopic data. I began with an initial T_{eff}, determined by either spectroscopic (preferred) or photometric methods. I calculated initial spectroscopic T_{eff}s using the Line-Depth Ratio (LDR) technique (Gray & Johanson 1991) with the selection of lines and the polynomial relations in Biazzo *et al.* (2007). I took the weighted mean of all T_{eff}s from all available line combinations, weighted by their sensitivity as stated in Biazzo *et al.* , Table 2. For spectra which did not contain sufficient data in the 6199-6275 Å region containing the lines in Biazzo *et al.* , I used photometry from the *SIMBAD Astronomical Database* to determine initial T_{eff}. The majority of the photometry contained in the SIMBAD database was taken from the Tycho (Høg *et al.* 2000) and 2MASS (Skrutskie *et al.* 2006) catalogs. I calculated photometric starting temperatures using the polynomial color-temperature-metallicity relation (ex: Ramírez & Meléndez 2005; Casagrande *et al.* 2010), and polynomial coefficients from Huang *et al.* (2015). The initial photometric T_{eff} for a given star was the mean of all Huang *et al.* temperatures, determined from all available SIMBAD color combinations, weighted by their error. Typically, the SIMBAD database contained B, V, R, J, H, and K photometry for a given star, which yields 5 separate color-temperature relations in Huang *et al.* . I corrected the photometry for the reddening values (for E(B-V)) as listed in Table 3.3.3, and adjust that value for other colors using the relations from Rieke & Lebofsky (1985). For both spectroscopic and photometric temperatures, when available color or line combinations resulted in more than 4 T_{eff} results, I omitted the highest and lowest T_{eff} before calculating a weighted mean.

I then adopt the gravity value (LogG), which fulfills the requirement that the titanium abundance, as measured using absorption features from the two ionization states (Ti I, Ti II),

is the same for both states. I then refine the $T_{\rm eff}$ values to fulfill the requirement that the measurement of iron abundance (Fe I) show no trend when compared to the excitation potential of each line. I iterate between the Ti I/Ti II "balance" and the Fe I "slope" processes until I arrive at a $T_{\rm eff}$ - LogG combination which fulfills both the balance and the slope requirements. This spectroscopic parameter determination process is well detailed in Takeda, Ohkubo & Sadakane (2002) and used by other automated processes like iSpec's "Equivalent Width Method" (Blanco-Cuaresma *et al.* 2014).

To incorporate the variances of my abundance measurements, I then performed an additional iterative process. I allow the Fe I "slope" parameter to vary by up to $\pm$0.05 dex/eV, and the Ti abundances to "match" as long as their 1-σ error bars overlap. This produces a range of acceptable $T_{\rm eff}$-$log\ g$ values, from which we selected the $T_{\rm eff}$-LogG pair which minimized the Fe I/Fe II abundance difference and the distance to a 1.5Gyr *PARSEC* (Bressan *et al.* 2012) isochrone at solar metallicity. I generated the *PARSEC* isochrones with the *CMD* tool `http://stev.oapd.inaf.it/cgi-bin/cmd`.

To determine the micro-turbulent velocity, I calculate the Ca abundance from selected absorption lines, using the previously-determined $T_{\rm eff}$ and LogG, for a range of ξ values. The final ξ value for each model is the one where the calculated Ca abundance variance is minimized. Figure 4.4a, from Lum & Boesgaard (2018) (hereafter: paper 1), illustrates the process for an example star in NGC 752. Note that while this is essentially the same process used in Reddy, Giridhar & Lambert (2012), I selected Ca to determine the minimum variance point, as opposed to the combination of Fe, V, Cr, and Ti in the prior work. The choice of Ca led to the ξ value with the least ambiguity, and Ca absorption lines were consistently available in both giant and main-sequence populations.

The resulting parameters were used to create atmospheric models using the *ATLAS9* grids (Castelli & Kurucz 2004). I used a linear interpolation algorithm between the existing grid points, based on our $T_{\rm eff}$, LogG, and ξ values. Since stellar metallicity is a parameter in determining the proper model, I used an iterative process to select the appropriate value for a given model. Starting with a solar metallicity, I calculated the Fe/H value for a given star, and then repeated the process, using the new metallicity, until I found a stable value. When a general metallicity value was required, notably in isochrone generation, I used a weighted average of the Fe Iand Fe IImeasurements from both giant and dwarf stars.

4.7.1 Absorption Line Evaluation

One of the secondary results of this project was to produce a method for selecting target lines, using objective criteria, from all available measured transitions within our spectral coverage range. Most comparable abundance studies use a substantially smaller set of "trusted" lines, adopted from previous, similar studies, and/or lines which have been reliable for the particular researcher in their prior work. In our own work (e.g. Boesgaard, Lum & Deliyannis (2015); Boesgaard, Roper & Lum (2013)), we have opted for a combination of the two selection processes.

For small-scale, high-resolution, high S/N studies, this process works well. However, for larger-scale and potentially lower S/N studies, a much larger sample of lines are needed. When compiling all the available line resources, including large online databases such as those run by VALD (Ryabchikova *et al.* 2015), and NIST (Kramida *et al.* 2014), a researcher can easily find a massive number of available lines. The challenge is identifying which to use.

For this project, I categorized each of the lines in the database by measuring the line in a high-resolution solar spectrum (Wallace *et al.* 2011), calculating the abundance of the respective element from that line width, and then comparing the result to the calculated solar abundances from Table 1 in Asplund *et al.* (2009). I then categorize the lines as "excellent", where the difference (Δ) in calculated abundance from that in Asplund *et al.* is less than 0.05 dex, "good" ($0.05 \leq \Delta \leq 0.10$), "fair" ($0.10 \leq \Delta \leq 0.20$), "poor" ($0.20 \leq \Delta \leq 0.40$), "bad/mismeasured" ($0.40 \leq \Delta \leq 1.00$), and "detected" ($\Delta \geq 1.00$). I assign each of these categories a numerical "Quality" score of 10 ("excellent"), 8, 6, 4, 2, or 1 ("detected"), and also use a score of 0 for lines which were either not detected in the solar spectrum, or which were measured using spectral synthesis.

Table 4.1 (from paper 1) shows abundances as calculated using each of the five above categories, plus the abundance as calculated using all measured lines. Using these calculations, I determined that the line choices for Pr II and Y I were not accurate enough to use for further analysis. Similarly, I was unable to measure the small number of S I and Eu II lines in the spectra to provide analysis of these elements.

Within the wavelength range of this study, the absorption lines for both C and N are weak or are blended with other nearby lines. While the automated line measurement process can evaluate blended lines of comparable strength, I chose to supplement the C and N abundance measures with additional measurements from synthesized spectra. I used MOOG's *synth* driver and line lists which combined elemental data from the VALD (Ryabchikova *et al.* 2015) database, with CN molecular data from Sneden *et al.* (2014). I discuss the specific regions and results for NGC 752 in in Section 5.6.2, and other synthesis results in the corresponding section for the other clusters.

4.7.2 Absorption Line Selection

Because I measured elemental abundances using a large numbers of lines in a particular spectrum, I have the option of using a subset of the lines. As stated earlier, when calculating elemental abundances, I only use lines which would lie on the linear portion of the "Curve of Growth" (CoG). Without going too deeply into the nature of spectroscopic absorption, the linear portion of the CoG occurs when the (Gaussian) core of an absorption feature expands in linear proportion to its width (on a log-log scale). This region provides the best resolution and most accurate measurement of elemental abundance from EQW measurements. While the exact point where this occurs varies with atmospheric conditions, I opt to use a hard cutoff of $\mathrm{Log}_{10}\,(\Delta\lambda/\lambda) < -4.80$ where $\Delta\lambda$ is the measured equivalent width, and λ is the central wavelength of the line.

The "linear CoG" selection eliminates a relatively small fraction of the measured lines, which in some cases (eg: Fe I, Ti I) still leaves dozens, if not hundreds of lines to choose from. While the immediate tendency might be to only select lines with the highest quality scores (as described above in Section 4.7.1), I also considered the effect of adding more samples to a given population in increasing the accuracy of a measurement. In general, adding more measurements of *equal quality* will result in a more accurate measurement of a given quantity. However, since I have an ad hoc evaluation of the quality of the line population, my process begins by selecting lines of the highest quality. Subsequent additions to my measurement subset are of lower quality. For many elements, I have a relatively large number of quality-rated lines, and feel the need to address the question of accuracy vs. precision.

Ion	Asplund (2009)	$\Delta \leq 0.05$	$\Delta \leq 0.10$	$\Delta \leq 0.20$	$\Delta \leq 0.40$	$\Delta \leq 1.00$	All lines
C I	8.43±0.05	8.44±0.02 (3)	8.42±0.04 (4)	8.41±0.09 (6)	8.47±0.13 (8)	8.55±0.25 (9)	8.55±0.25 (9)
O I	8.69±0.05	8.69±0.02 (2)	8.72±0.05 (3)	8.77±0.06 (6)	8.77±0.06 (6)	8.84±0.18 (7)	8.84±0.18 (7)
Na I	6.24±0.04	6.26±0.01 (2)	6.26±0.01 (2)	6.27±0.08 (5)	6.23±0.12 (6)	6.36±0.32 (7)	6.36±0.32 (7)
Mg I	7.60±0.04	7.62±0.02 (6)	7.61±0.05 (10)	7.60±0.07 (12)	7.60±0.07 (12)	7.48±0.27 (15)	7.20±0.91 (17)
Al I	6.45±0.03	6.42±0.00 (1)	6.41±0.01 (2)	6.35±0.06 (4)	6.31±0.08 (6)	6.31±0.08 (6)	6.31±0.08 (6)
Si I	7.51±0.03	7.53±0.03 (8)	7.52±0.05 (16)	7.57±0.10 (33)	7.59±0.11 (37)	7.62±0.15 (40)	7.71±0.41 (42)
Si II	7.51±0.03	—	—	7.70±0.00 (2)	7.70±0.00 (2)	7.70±0.00 (2)	7.70±0.00 (2)
S I	7.12±0.03	—	—	7.27±0.00 (1)	7.37±0.10 (2)	7.37±0.10 (2)	7.37±0.10 (2)
Ca I	6.34±0.04	6.31±0.01 (3)	6.32±0.05 (5)	6.32±0.14 (14)	6.23±0.21 (29)	6.15±0.27 (35)	6.15±0.27 (35)
Sc I	3.15±0.04	—	3.22±0.00 (1)	3.19±0.13 (5)	3.18±0.19 (7)	3.18±0.19 (7)	3.18±0.19 (7)
Sc II	3.15±0.04	3.14±0.01 (4)	3.15±0.05 (7)	3.15±0.09 (11)	3.16±0.15 (13)	3.19±0.18 (14)	3.19±0.18 (14)
Ti I	4.95±0.05	4.95±0.03 (60)	4.94±0.05 (103)	4.94±0.10 (166)	4.97±0.19 (255)	5.09±0.36 (344)	5.55±0.96 (471)
Ti II	4.95±0.05	4.94±0.03 (9)	4.96±0.06 (16)	4.97±0.10 (25)	4.95±0.21 (41)	5.03±0.36 (54)	5.13±0.69 (59)
V I	3.93±0.08	3.92±0.02 (24)	3.94±0.05 (40)	3.95±0.09 (58)	3.94±0.12 (64)	3.96±0.15 (66)	3.99±0.28 (67)
V II	3.93±0.08	3.92±0.00 (1)	3.88±0.04 (2)	3.86±0.04 (3)	3.79±0.13 (4)	3.79±0.13 (4)	3.79±0.13 (4)
Cr I	5.64±0.04	5.63±0.03 (20)	5.62±0.06 (46)	5.63±0.11 (78)	5.63±0.16 (100)	5.68±0.33 (130)	5.79±0.58 (141)
Cr II	5.64±0.04	5.63±0.03 (2)	5.64±0.06 (4)	5.67±0.10 (7)	5.62±0.20 (12)	5.58±0.29 (15)	5.58±0.29 (15)
Mn I	5.43±0.05	5.43±0.02 (5)	5.41±0.06 (11)	5.36±0.08 (20)	5.34±0.09 (23)	5.34±0.09 (23)	5.41±0.34 (24)
Fe I	7.50±0.04	7.49±0.03 (49)	7.49±0.05 (79)	7.47±0.10 (148)	7.39±0.19 (265)	7.32±0.30 (355)	7.32±0.39 (368)
Fe II	7.50±0.04	7.50±0.04 (12)	7.49±0.05 (22)	7.48±0.09 (35)	7.43±0.15 (46)	7.37±0.27 (56)	7.34±0.32 (58)
Co I	4.99±0.07	4.99±0.03 (19)	4.96±0.05 (33)	4.96±0.09 (50)	4.99±0.17 (66)	5.08±0.37 (92)	5.69±1.10 (130)
Ni I	6.22±0.04	6.22±0.03 (29)	6.23±0.06 (62)	6.24±0.10 (97)	6.23±0.16 (125)	6.22±0.29 (150)	6.29±0.51 (158)
Cu I	4.19±0.04	4.16±0.00 (1)	4.18±0.08 (3)	4.18±0.08 (3)	4.23±0.12 (4)	4.31±0.19 (5)	4.31±0.19 (5)
Zn I	4.56±0.05	—	4.65±0.00 (2)	4.67±0.03 (3)	4.67±0.03 (3)	4.67±0.03 (3)	4.67±0.03 (3)
Y I	2.21±0.05	—	—	—	2.43±0.00 (1)	2.63±0.21 (2)	2.63±0.21 (2)
Y II	2.21±0.05	2.20±0.01 (2)	2.17±0.04 (3)	2.26±0.09 (7)	2.34±0.19 (14)	2.25±0.29 (16)	2.25±0.29 (16)
Zr I	2.58±0.04	2.55±0.00 (1)	2.59±0.05 (2)	2.64±0.11 (6)	2.66±0.11 (7)	2.66±0.11 (7)	2.66±0.11 (7)
Zr II	2.58±0.04	2.56±0.00 (1)	2.56±0.00 (1)	2.56±0.00 (1)	2.56±0.00 (1)	2.56±0.00 (1)	2.56±0.00 (1)
Ba II	2.18±0.09	—	—	2.03±0.04 (2)	2.15±0.17 (3)	2.04±0.24 (4)	1.84±0.46 (5)
Ce II	1.58±0.04	1.61±0.01 (3)	1.59±0.06 (8)	1.61±0.08 (9)	1.65±0.15 (14)	1.68±0.17 (15)	1.93±1.00 (16)
Pr II	0.72±0.04	—	—	—	—	1.60±0.00 (1)	1.60±0.00 (1)
Nd II	1.42±0.04	1.42±0.02 (8)	1.42±0.03 (10)	1.45±0.07 (12)	1.47±0.11 (13)	1.52±0.20 (14)	1.52±0.20 (14)
Sm II	0.96±0.04	0.96±0.00 (1)	1.02±0.04 (4)	1.02±0.04 (4)	1.09±0.11 (6)	1.15±0.17 (7)	1.15±0.17 (7)
Eu II	0.52±0.04	—	0.45±0.00 (1)	0.45±0.00 (1)	0.45±0.00 (1)	0.71±0.26 (2)	0.71±0.26 (2)

TABLE 4.1: Solar absorption line quality evaluation

By hand-selecting a small number of lines, a researcher can guarantee a precise result - simply put, they can select for small error bars. Conversely, a set of lines exists which would provide the most accurate measurement of the abundance of a given element in a given star. Unfortunately, we cannot know the latter set, without knowing the actual abundance(s) from another independent measure. For the purposes of this study, I am interested in the most accurate measurement of atmospheric abundances. However, I also wish to have at least a "reasonable" level of precision. So, in order to answer the question of accuracy vs. precision, I created a Monte-Carlo simulation of the line measurement process.

The simulation starts with an assumption of an "actual" elemental abundance of $[X/Fe] = 0.00$. It randomly generated a set of between 1 and 250 lines, distributed among the 6 "quality" categories as indicated by the population in the solar measurements (Table 4.1) - Approximately 16% in the top ("10") category, 13% in the next ("8"), and 20%, 21%, 18%, and 12% in the remaining four categories, respectively. For each line, it generated an experimental abundance measurement, using a randomized Gaussian distribution around the solar value. The standard deviation (σ) used was the average σ of all the lines in the corresponding category in the solar measurements, added in quadrature with an atmospheric term of 0.10, representative of a typical atmospheric parameter error term (see Table 5.3 for NGC 752). By comparing the experimental abundance, calculated as the mean of all lines of a given quality or better, with the actual value of 0.00, it determined the quality-number relationship used herein.

The simulation ran 10,000 trials, using N=$1, 3, 5, 10, 20, 50, 100$, and 250 lines to measure the simulated abundance of the simulated element. It then evaluated both the accuracy and precision of abundances calculated from lines of varying "quality" ratings. Unsurprisingly, the most precise abundances were calculated when only the lines of the highest available quality ranking were used. However, as the number of available lines for a given element increase, I find that including more lines of lower quality does increase the accuracy of the measurement.

In summary, the simulation found, for elements where it measure 4 or fewer lines, using only the lines in the highest available quality category provided the most accurate abundance measurement. For example, if it measured 3 lines with quality scores of 10, 8, and 6, it would calculate the abundance, using only the single line of the highest quality (10). Should none of the measured lines for a given element have the highest quality rating, it would take the abundance as measured by lines in the next lower category, reducing the acceptable quality category until at least one line was used.

In the case where it measured between 5 and 10 lines for an element, it would not discriminate between lines of the highest two categories (10, 8), and would use all available lines in these two quality categories. As in the cases with small numbers of lines, should none of the measured lines rank in the top two quality categories, it calculates the elemental abundance using the highest quality lines available from the measured set.

For $N \geq 20$, the most accurate abundance values are achieved by using all measured lines in the top 3 categories. In all spectra for this project, when 20 or more lines were measured for a given element, a large fraction (generally the majority) of lines fell within the top three quality categories. Therefore the abundances calculated for large-N elements always used the "best" lines.

4.8 Elemental Abundance Calculation

Earlier, I have briefly mentioned the use of *MOOG* (Sneden 1973; Sneden *et al.* 2012) in transforming EQW measures into elemental abundances. To elaborate a bit further on the process: independent abundance measurements were obtained for each measured EQW through use of MOOG's *synth*, *blends*, and *abfind* "drivers". Generally, MOOG's operation was applied as a "black box" translation from EQW to abundance, but the general functioning of MOOG is to create a simulated spectrum, using a passed line list, at the base of the passed model atmosphere, then evolve that spectrum through each of the layers of a passed model atmosphere file. The process used by MOOG is a balance between accurate simulation and operational efficiency (memory and processing time). In this respect, a number of sacrifices are made (e.g..: No magnetic field interaction), but the rapid execution of a MOOG "run" allows the results to be "good enough".

While this project's analysis spanned several *MOOG* version upgrades, between 2014 and 2017, there is no significant difference in results between the earliest (July 2014) and latest (February 2017) versions. Note that I made superficial and cosmetic modifications to the standard MOOG program and make files to allow compilation under the *GNU Compiler Collection's GFortran* (Free Software Foundation, `https://gcc.gnu.org/`) compiler, and removed status messages to allow the *MOOGSILENT* function to run truly silently.

MOOG expects two files as input, a line list, and a model file. A third command file is also required for the automated operation used by this software. The line list file is a list of the measured EQWs and the physical parameters (e.g.: excitation potential, $\text{Log}g_f$) of the associated lines. The model file is a text file with a list of conditions (pressure, temperature, etc.) at ascending levels of a stellar atmosphere. The command file is as described in the MOOG documentation[8]. My software then parses the output files from MOOG, and creates the desired output format (table, plots, etc. — see next section).

4.9 Output/Assessment — Tables and Plots

The output from the MOOG package is a text file, containing The basic output from abundance evaluations of cluster stars is a .tex file, suitable for inclusion in LaTeXdocuments, as in Table 5.4. I also found the need to create a series of plots to assist with atmospheric parameter evaluation, and to clearly represent composite cluster abundances. All plots were created using the Python "matplotlib" library.

4.9.1 Atmospheric Parameter Determination

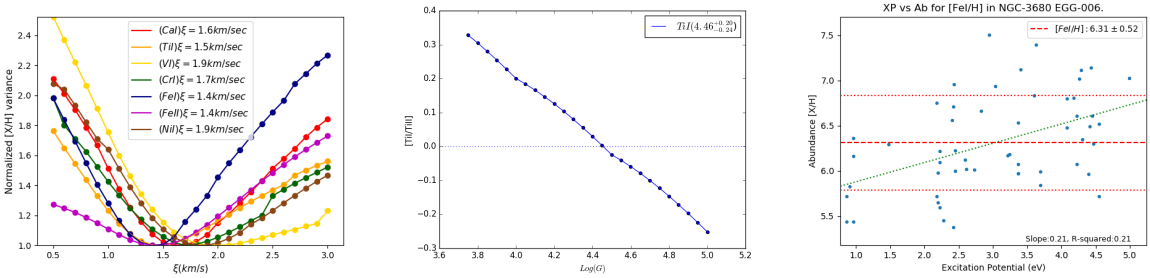
One of the benefits of automating the abundance determination process is that I can run it multiple times, varying atmospheric parameters with each run, and then compare results. In Section 4.7, I described the use of Ca abundance variance minimization. Figure 4.4a is a graphic comparison of the potential elements used in V_{turb} determination. The value I selected is determined by the Ca "Δ", but I also show elements used by Reddy, Giridhar & Lambert.

Figure 4.4b demonstrates the process for determining LogG, as described in Section 4.7, above. The plot is created by holding T_{eff} and V_{turb} constant, and varying LogG. The point

[8] `https://www.as.utexas.edu/~chris/codes/WRITEMOOG.ps`

at which the abundance "Δ" between the un-ionized and singly-ionized species of Ti is zero is selected as the base parameter value, which is then adjusted for an isochrone "prior".

$T_{\rm eff}$ is similarly graphically determined. Figure 4.4c plots Fe abundance against the excitation potential of individual lines in the star Egg-6, in NGC 3680. At the "correct" $T_{\rm eff}$ for a given star, there should be no trend between Fe abundance and excitation potential (XP). The green dotted line represents a linear fit to the Fe abundance vs. excitation potential of this star at the "Guessed" temperature. The $T_{\rm eff}$ used in making this plot was (intentionally) set 400K below the $T_{\rm eff}$ selected for abundance calculation, in order to show a slope in the Ab(Fe)[9] vs. XP relationship.

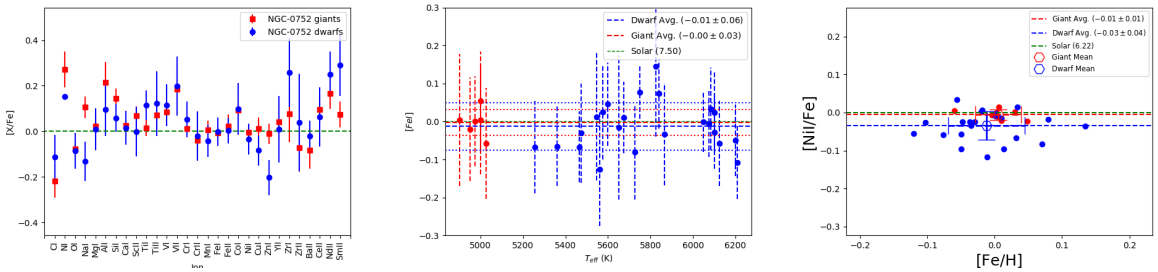


(A) ξ in Pla-1089 (B) Log G in Pla-300 (C) $T_{\rm eff}$ in Egg-6

FIGURE 4.4: Illustration of the use of multiple iterations of abundance calculations in atmospheric parameter determination. The $T_{\rm eff}$ value for the (NGC 3680) star Egg-6 in Figure 4.4c (right) is intentionally set 400K low, in order to demonstrate a trend in Ab(Fe)[9] vs. excitation potential. This trend disappears when the "correct" $T_{\rm eff}$ is used.

[9]The abundance notation of Ab(X) for a given element "X" is a shorthand for $log_{10}(N(X)/N(H)) + 12.00$ where $log_{10}N(H)$ is set at 12.00.

4.9.2 Cluster Abundance Visualization



(c) T_{eff} in Egg-6

FIGURE 4.5: Examples of three graphical representations of cluster abundances. Each point in Figure 4.5a (left) represents the mean abundance (with associated error bar) for one element for all dwarfs (blue) or giants (red) in a cluster. Figure 4.5b (center) plots the abundance for one element, Fe in this case, for each star (in one cluster) against T_{eff}. The final example (right) plots each star's abundance, and the respective population mean (and uncertainty) for a given element, against that star/population's [Fe/H].

I also wrote some short functions to display abundance results in a graphical format for quick visual summary. The data for these plots is the same as used for the abundance table discussed in Section 4.9. Three example element displays are shown in Figure 4.5.

4.9.3 Other Plots

During the development process, I also experimented with other methods of element/abundance displays, detection sensitivity, curve-of-growth selection, etc. I also wrote code to create and display the Monte Carlo results discussed in Sections 4.5.1 and 4.7.2. While none of these outputs are relevant in my final analysis, the code (and some of the plots) are still present in the project repository. I will re-assess the relevance of this code during the next maintenance release of the code base.

4.10 Other Scripts

Also included in my software package is the analysis code, and display routines used for cluster membership analysis in Chapter 3. While these are mostly contained in the "MembershipDetermination" module, the routines are not stand-alone, and require constants and utilities from the overall project. The online ID cross-referencing is also performed using another module.

One of the benefits of the Python language is that it is a scripting language. As such, it is extremely easy to "fire off" a quick script to perform simple tasks or data inquiries. Unfortunately, the ease and convenience of writing such scripts means that proper design documentation is sacrificed in the quest for convenience. In a properly designed project, none of these "one-offs" would exist. In the interest of removing these from the released software, I have tried to limit these scripts to the "test" temporary project. However, I did not start this

practice until later in the project's development, so there may still be a few of these rogue script files floating around the project. Likewise, I have written single-use scripts for entering new lines and references into the absorption line database. These scripts are a sloppy way of maintaining the database, and need to have the more refined interface and error handling that would come with being a full sub-module of the project.

Chapter 5

NGC 752: A Case Study

My analysis process is fairly complex, and while it can be determined from the software package descriptions in the prior chapter, I feel it best to describe this process through a single cluster walk-through, or "case study." I am also following the same procedure in publishing my results. This case study of NGC 752 was the first to be published (paper 1), mainly for the purpose of getting the process out to the scientific community.

5.1 NGC 752

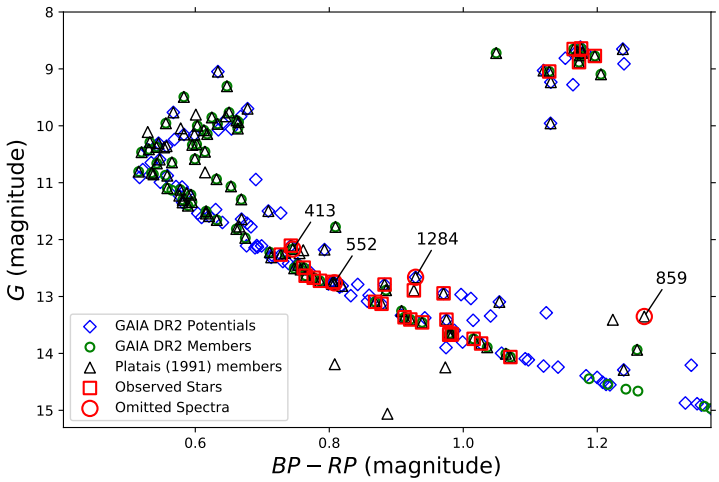


FIGURE 5.1: Color-magnitude diagram for NGC 752 (from: Lum & Boesgaard 2018). Spectra which were available from the Keck Observatory Archive are noted in red. Spectra which were omitted for various reasons (see text) are numbered with their Platais (1991) IDs.

I selected NGC 752 as the first cluster in my study due to a number of factors. It is a member a relatively small number of “solar twin” clusters, having a metallicity approximately solar (see Section 2.5). For calibration purposes, using near-solar metallicity targets permits me to use the Sun as a reference spectrum to confirm laboratory-determined absorption line parameters with observational measurements. NGC 752 is relatively old open cluster, with an age between 1.4 and 1.8 Gyr. However, even at this age, the MSTO stars should still be massive enough to have used the CNO cycle as the primary (fusion) energy generation process in their

cores. Therefore, NGC 752 is a prime target for detecting evidence of CNO processing in the atmospheres with "dredged up" elements. Finally, as a Northern Hemisphere cluster with a declination of $+38^{\circ}$, all of the high-resolution spectra for this cluster come from a single instrument, the HIRES spectrograph on the Wm. Keck I telescope. By limiting this first cluster to one instrument/archive, I was able to reduce the amount of processing needed for the spectra.

Figure 5.1 places each of my "observed"/target stars on a color-magnitude diagram, using the procedure detailed in Chapter 3. Also noted on the figure are the stars denoted as members by Platais (1991), in red. I have also included stars from the GAIA dataset which appear to be members, or potential members, based on analysis of position, proper motion, and parallax as per Chapter 3.

At NGC 752's age, the main-sequence turn off stars are F-type stars. These stars lie on the blue side of the Li "gap" (Boesgaard & Tripicco 1986), implying that sufficient mixing occurs during the main sequence lifetime of the star to have carried surface material to a depth where Li can be destroyed by (p,α) reactions. Hobbs & Pilachowski documented both the Li destruction in NGC 752 giants (Pilachowski, Saha & Hobbs 1988) and the presence of the "Li gap" in the F-stars (Hobbs & Pilachowski 1986; Pilachowski & Hobbs 1988). Deep mixing mechanisms may also affect atmospheric elemental abundances, as measured at the surface, after the first dredge up portion of the red giant branch.

5.2 Stellar Spectra

Table 5.1 (from paper 1) lists the J2000 coordinates (RA/Dec), spectral wavelength range, and the signal-to-noise for the composite spectra of each star in our survey. We chose the Platais (1991) identifiers for our reference, but have included cross-references for Heinemann (1926) and Rohlfs & Vanysek (1962) (WebDA) as reported by the SIMBAD astronomical database (Wenger *et al.* 2000). The wavelength ranges cited reflect the minimum and maximum values for each spectrum. The spectral coverage for a given range is generally not complete, due to inter-order gaps and gaps between coverage of individual observation configurations.

All NGC 752 spectra were obtained from the Keck Observatory Archive (KOA) and were measured using the HIRES spectrograph, with a resolution of ~48,000, as calculated from the instrument setup parameters (KOA "SPECRES" keyword). We utilize observations from both the "original" (prior to August 2004) and "upgraded" (post-2004) HIRES. The upgraded detector added two CCDs, for a total of three, increasing the wavelength coverage while reducing pixel size from 24μ to 15μ .The difference between the two detectors is significant, and manifests in our data as a difference in the signal-to-noise ratio with respect to exposure time, and increase in wavelength coverage for a given spectrum.

TABLE 5.1: NGC 752 Archive Observations

Reference Number			RA (J2000)	Dec (J2000)	Wavelength	S/N		Date of	Observer
Pla.	Hen.	Rohlfs	(hh:mm:ss)	(Deg:mm:ss)	Range (Å)	6200 Å	7800 Å	Obs.	Inits.
				Dwarfs					
300	—	487	01:55:27	+38:08:32	4350-6860	60	—	2000-10-08	BJ
361	29	310	01:55:44	+37:54:31	4350-6860	50	—	2000-10-08	BJ
391	38	313	01:55:53	+37:49:26	6240-8680	—	70	2000-10-09	BJ
413‡	48	294	01:55:59	+37:40:49	6240-8680	—	70	2000-10-09	BJ
429	54	288	01:56:04	+37:36:42	6240-8680	—	70	2000-10-09	BJ
520	80	290	01:56:23	+37:38:14	5650-8090	150	130	2003-11-02	AB
552‡	87	280	01:56:32	+37:34:33	6240-8680	—	60	2000-10-09	BJ
575	94	303	01:56:36	+37:45:35	6240-8680	—	40	2000-10-09	BJ
699*	144	437	01:57:04	+38:07:20	4350-6860	70	—	2000-10-08	BJ
					5650-8090	150	130	2003-11-02	AB
701	146	260	01:57:06	+37:50:43	6240-8680	—	60	2000-10-09	BJ
					5650-8090	90	80	2003-11-02	AB
					4000-8490	90	130	2009-08-27	JC
					4000-8490	90	120	2009-08-27	JC
786	183	270	01:57:22	+37:38:21	4350-6860	70	—	2000-10-09	BJ
					4000-8490	90	130	2009-08-27	JC
					4000-8490	80	120	2009-08-27	JC
790	185	—	01:57:24	+37:52:12	5650-8090	160	140	2003-11-02	AB
791	184	434	01:57:24	+38:06:10	5650-8090	150	130	2003-11-02	AB
828	199	266	01:57:34	+37:42:01	6240-8680	70	—	2000-10-09	BJ
859**	207	252	01:57:38	+37:49:44	6240-8680	70	—	2000-10-09	BJ
					4000-8490	110	140	2009-08-27	JC
					4000-8490	70	100	2009-08-27	JC
864	211	430	01:57:39	+38:08:39	6240-8680	—	40	2000-10-09	BJ
					5650-8090	150	130	2003-11-02	AB
889	216	427	01:57:45	+38:11:07	6240-8680	—	30	2000-10-09	BJ
					5650-8090	150	130	2003-11-02	AB
921	229	103	01:57:52	+37:27:46	5650-8090	160	140	2003-11-02	AB
964	244	388	01:58:03	+38:02:30	5650-8090	130	120	2003-11-02	AB
993	256	101	01:58:09	+37:28:36	6240-8680	—	60	2000-10-09	BJ
999	—	107	01:58:11	+37:23:53	6240-8680	—	70	2000-10-09	BJ
1012	—	118	01:58:13	+37:15:20	5650-8090	150	120	2003-11-02	AB
1017	265	240	01:58:16	+37:33:26	6240-8680	—	70	2000-10-09	BJ
1107	298	232	01:58:34	+37:40:19	6240-8680	—	70	2000-10-09	BJ
1270	—	—	01:59:19	+37:49:50	6240-8680	—	70	2000-10-09	BJ
1284‡	—	—	01:59:20	+37:23:23	6240-8680	—	70	2000-10-09	BJ
					4000-8490	100	140	2009-08-27	JC
					4000-8490	110	150	2009-08-27	JC
1365	—	187	01:59:47	+37:49:48	4350-6860	35	—	2000-10-08	BJ
				Giants					
350	24	25	01:55:40	+37:52:28	3360-8100	55	70	2008-06-20	GM
					3360-8100	160	220	2008-06-22†	GM
356	20	27	01:55:43	+37:37:39	3360-8100	55	60	2008-06-20†	GM
					3360-8100	160	220	2008-06-22	GM
506	29	75	01:56:19	+37:58:02	4690-9140	290	400	2009-10-08	SS
687	—	137	01:57:04	+38:07:57	3360-8100	50	70	2008-06-20	GM
					3360-8100	160	220	2008-06-22	GM
					4690-9140	350	450	2009-10-08	SS
1089	67	295	01:58:30	+37:51:37	3360-8100	50	70	2008-06-20	GM
					4690-9140	290	400	2009-10-08	SS
1172	65	311	01:58:53	+37:48:57	3360-8100	50	80	2008-06-20	GM

* - Possible single-line binary
† - Possible mis-aligned order fit (see text)
‡ - Spectroscopic Binary (unused)
** - Parallax differential (unused)

AB- Boesgaard
JC- Cohen
BJ- Jones
GM- Marcy
SS- Schuler

The minimum measured S/N ratios of 40-60 for the dwarf sample are similar to those of comparable studies - e.g.: Maderak *et al.* (2013) had a minimum S/N of 60. The S/N ratio of the giant star spectra in this study are significantly higher than those of other studies, topping out at well over 400. The S/N range in the Böcek Topcu *et al.* (2015) giant study was 140-290, while Reddy, Giridhar & Lambert (2012) had S/N ratios in the 100-120 range.

The processing of the downloaded spectra followed the process as detailed in Section4.4. As an additional quality control process, I performed a visual inspection of the individual orders, as provided in the Keck Observatory Archives image previews. I opted to exclude two orders (out of several hundred) which showed evidence of reduction or observational faults. In both cases, the missing orders were conveniently duplicated in other spectra, without the flaws.

During my analysis process, I also elected to exclude spectra for four stars from the analysis process. Although high resolution and high S/N spectra were available for PLA-413, PLA-552, PLA-859, and PLA-1284, they were excluded for the following reasons: For PLA-859, the parallax data from GAIA DR2 (0.28 ± 0.04mas) places it beyond the cluster (mean parallax from GAIA DR2 data of 2.2 ± 0.1, from Chapter 3). For PLA-1284, the automated process for measuring absorption lines in this (SB2) binary system did not properly account for the shifted components. After analysis, I found that the spectroscopic binary, PLA-552 displayed evidence of line broadening, which affected abundance calculations (particularly of Ti and Cr). While PLA-413 is not cataloged as a spectroscopic binary, it lies above NGC 752's main sequence, near other binaries, and displays identical broadened metal lines. For this reason, I also excluded both PLA-552 and PLA-413 from additional analysis.

5.3 Equivalent Width Measures

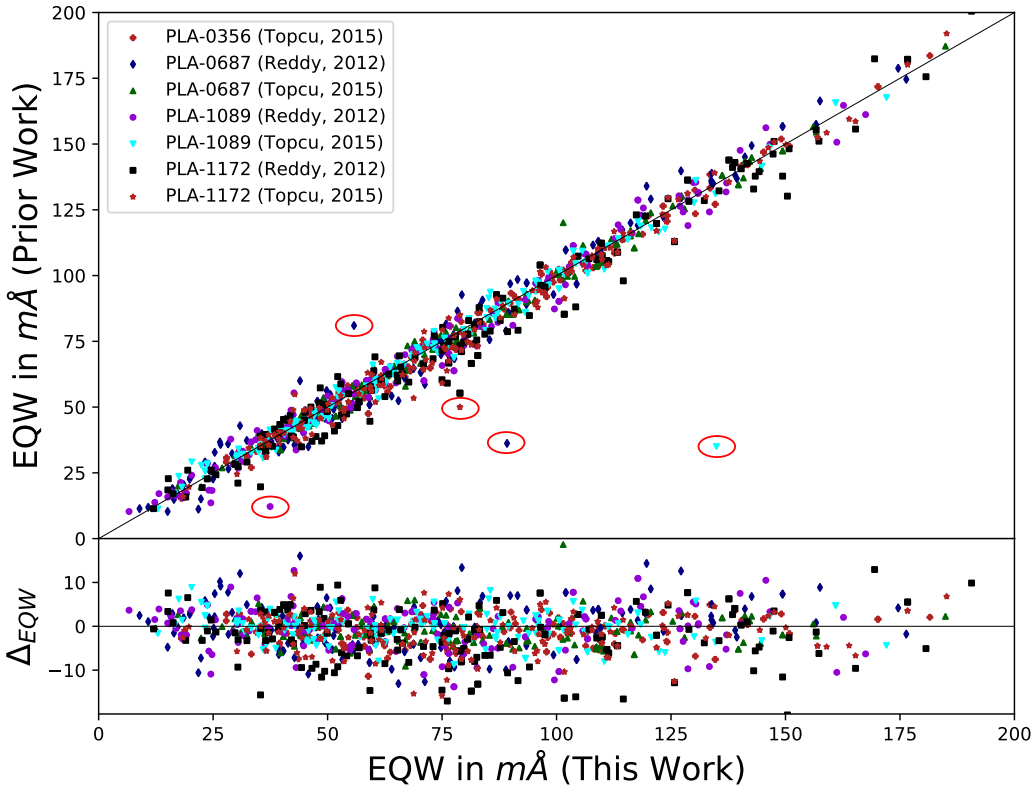


FIGURE 5.2: Comparison of 1211 Equivalent Width (EQW) measurements between Reddy, Giridhar & Lambert (2012), Böcek Topcu *et al.* (2015), and this work for common giant stars. The circled outliers are discussed in the text. (From Lum & Boesgaard (2018))

I ran the corrected and stacked spectra through the equivalent width measurement process as detailed in Chapter 4. The minimum acceptable EQW measurement for each spectra, calculated as a function of dispersion and signal-to-noise ratio, is listed in the final column of Table 5.2.

As NGC 752 was the pilot cluster, testing the overall measurement and abundance calculation process, I felt it was necessary to compare the intermediate results of the process with similar results from literature, when possible. For the EQW measurement process, I found that both Reddy, Giridhar & Lambert (2012) and Böcek Topcu *et al.* (2015) had published EQW measurements for stars which were common with my data set. In both studies, the method used was to manually fit a Gaussian curve to each feature, using STSDAS's IRAF *splot* function. This is also the method I have used in prior work (Boesgaard, Roper & Lum 2013; Boesgaard, Lum & Deliyannis 2015), and is the primary method I am aiming to replace with this automated process. Figure 5.2 plots measurements of the same line in the same star

between different works in the upper portion of the graph, and the "Delta" between measurements of a given line, below. On the whole, I found no systematic different between measurement methods ($\Delta_{EQW} = 1.2 \pm 5.9$ mÅ). I then assessed the five largest outliers to see if there was a major fault in my analysis, which would affect large-scale measurements.

In order of increasing EQW, as measured by the automated process: 1) The EQW of the Ce II line at 5274.21 Å was 37.4 mÅ while Reddy, Giridhar & Lambert (2012) record a measurement of 12.2 mÅ. I manually re-measured the line and did not find a significantly different measurement from the automated process, and suspect that the difference is due to either a mis-measurement or a typographical error with Reddy, Giridhar & Lambert, as their EQW measures for the same line in their other giants are similar to mine. In PLA-687 they measure 39.9 mÅ vs. my measurement of 35.6 mÅand in PLA-1172, 39.0 mÅ (42.1 mÅ from my process). 2) For the Ti II line at 4911.19 Å, I determined that my measurement (of 55.8 mÅ) was affected by a form of signal clipping, probably introduced during the cosmic ray removal process, which caused the measurement process to underestimate the line width. Since this is an unusual occurrence, I elected to retain my measurement as taken, and account for it as an expected random variation. 3) I elected to treat the Co I line at 5212.69 Å (my EQW=78.9 mÅ) in a similar fashion. This line is blended with a Ti I line at 5212.28 Å, and my automated process fits the Co I and Ti I lines as a two-component blend. Manual fitting did not yield a significant difference. The final two selected outliers; 4) a Zn I line at 4722.15 Å in PLA-687 (my EQW=89.2 mÅ, Reddy, Giridhar & Lambert's EQW=36.2 mÅ) and 5) a Fe II line at 6246.32 Å in PLA-1089 (my EQW=134.9 mÅ, Böcek Topcu *et al.* 's EQW=35.0 mÅ), also fall into the "probable error in the prior work" (likely typographical or transposition error) category, as those measurements are also outliers in their respective work, relative to measurements of the same lines in other stars.

5.4 Atmospheric Parameters

Following the processes as detailed in Chapter 4, I determined the physical parameters for both the dwarfs and giants in NGC 752. The final parameters for each star's $T_{\rm eff}$, LogG, and micro-turbulent velocity (ξ) are listed in Table 5.2.

TABLE 5.2: NGC 752 Atmospheric Parameters

PLA #	T_{eff} (K)	log g	ξ (km s^{-1})	*min.EQW* (mÅ)
		Dwarfs		
300	5650	4.52	2.1	12.5
361	5545	4.55	1.8	8.3
391	5465	4.56	1.8	6.3
429	5255	4.57	1.8	6.5
520	6100	4.44	1.6	2.6
575	5600	4.52	1.8	8.8
699	5865	4.50	1.6	2.7
701	5725	4.52	1.7	1.5*
786	5560	4.55	2.0	1.6*
790	6210	4.44	2.1	3.0
791	6100	4.42	1.8	3.0
828	5475	4.53	1.9	6.3
864	6085	4.43	1.7	3.0
889	6125	4.46	1.7	3.0
921	6075	4.46	1.6	2.7
964	6050	4.43	1.6	3.6
993	5575	4.54	1.4	6.5
999	5750	4.50	1.4	5.8
1012	6200	4.40	1.7	3.3
1017	5840	4.47	1.6	6.0
1107	5675	4.52	1.7	5.6
1270	5360	4.55	1.8	5.6
1365	5825	4.46	1.4	11.2
		Giants		
350	5000	2.69	1.6	1.9*
356	4950	2.68	1.5	1.9*
506	4975	2.70	1.5	1.1*
687	5000	2.74	1.5	0.7*
1089	5025	2.66	1.5	1.0*
1172	4900	2.71	1.6	4.0
*-Adjusted to minimum of 2.0 mÅ				

I considered the atmospheric parameters as another intermediate data product which could be compared to prior work. Figure 5.3 (from paper 1) does just that. Generally speaking, my parameters are within errors of their counterparts as measured by others (Sestito, Randich & Pallavicini 2004; Maderak *et al.* 2013; Reddy, Giridhar & Lambert 2013; Böcek Topcu *et al.* 2015; Castro *et al.* 2016). Different methods are represented in the diagram. Maderak *et al.* and Castro *et al.* selected points directly from isochrones. While Sestito, Randich & Pallavicini simply chose a log g value of 4.50 for all of their dwarf stars. My $T_{\rm eff}$ - LogG points differ slightly from the isochrone line, most likely due to the uncertainty in abundance measurements used in the parameter determination process. However, by allowing my points to differ from the isochrone, I also allow parameter variation due to factors not accounted for in the isochrone models, such as magnetic field and rotational effects. By only using

the isochrone values as a prior (in a quasi-Baysean sense) consideration, my parameter determination process remains observationally based, as opposed to theoretical- or model-based.

Based on comparison between my selected parameter values, those from prior works, isochrone models, and the resolution of my iterative parameter determination process, I selected conservative uncertainty values for the parameters of $T_{\rm eff} \pm 100K$, LogG $\pm$ 0.05, and $V_{\rm turb}(\xi) \pm 0.10$ km s^{-1}.

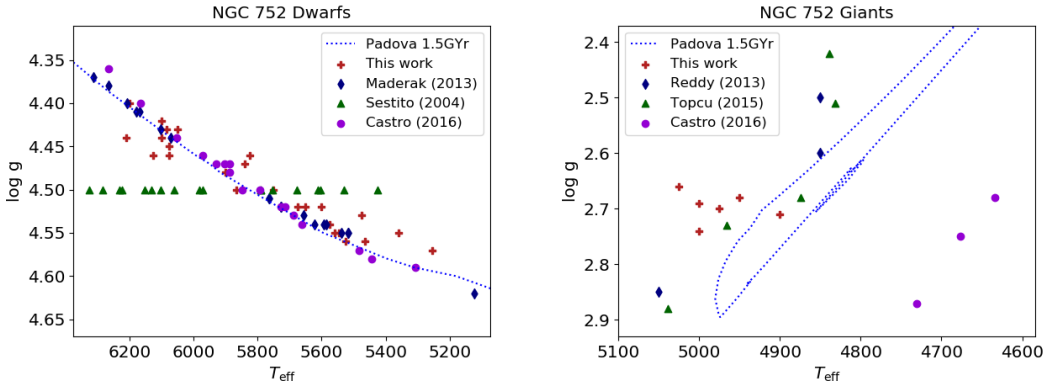


FIGURE 5.3: Graphical comparison of parameters used between stars in common between, Sestito, Randich & Pallavicini (2004); Maderak *et al.* (2013); Reddy, Giridhar & Lambert (2013); Böcek Topcu *et al.* (2015) and Castro *et al.* (2016). Note that the close adherence of the Maderak *et al.* and Castro *et al.* points to the dwarf isochrone is due to the fact that they selected points directly along the isochrone line. Likewise, Sestito, Randich & Pallavicini simply chose a log g value of 4.50 for all of their dwarf stars.

5.5 Abundance Calculation Uncertainty

There are two (major) sources of uncertainty represented in my abundance calculations. In calculating an element's abundance for a given star, I take the mean of all of the abundances selected according to the process in Section 4.7.2. A natural uncertainty resulting from this operation is the variance between individual measurements. In Table 5.4, abundance uncertainties due to line-to-line measurement variance is listed in the "σ_{var} column.

As noted in the previous section, all three atmospheric parameters have an uncertainty. Variation of atmospheric parameters leads to variance in elemental abundance calculations for a given star. In order to properly characterize the uncertainty in individual star abundance measurements, caused by varying model atmosphere parameters, I re-ran the full abundance determination process, varying each of the three parameters, as shown in Table 5.3. I then added the largest error of each parameter (either the "+", or the "-" error) in quadrature, to produce an atmospheric parameter error. Abundance uncertainty, due to uncertainty in atmospheric parameters, is designated as σ_{atm} in Table 5.4. For the purposes of the remaining clusters in this project, the abundance uncertainties determined from the NGC 752 stars' parameter variance, are considered "typical", so I adopted the σ_{atm} values from NGC 752 for all clusters.

5.6 NGC 752 Abundance Results

My abundance results are shown in Table 5.4, which contains the cluster averages, with dwarf and giants averaged separately, along with the standard deviation of the individual star measurements,

TABLE 5.3: NGC 752 Atmospheric Errors

Parameter	C I	N I	O I	Na I	Mg I	Al I	Si I	Si II	Ca I	Sc I	Sc II	Ti I	Ti II	V I	V II	Cr I
Dwarfs																
$T_{\rm eff}+100K$	-0.01	-0.08	-0.10	0.04	0.03	0.05	0.01	-0.10	0.07	0.10	0.00	0.09	-0.01	0.10	-0.01	0.07
$T_{\rm eff}-100K$	0.01	0.06	0.10	-0.04	-0.03	-0.05	-0.00	0.10	-0.07	-0.10	-0.00	-0.10	0.02	-0.11	0.01	-0.08
log g+0.02	0.00	0.01	0.01	-0.00	-0.00	0.00	0.00	0.01	-0.00	0.01	0.01	0.00	0.01	0.00	0.01	-0.00
log g-0.02	-0.00	-0.01	-0.01	0.00	0.00	-0.00	-0.00	-0.01	0.00	0.00	-0.01	-0.00	-0.01	-0.00	-0.01	-0.00
$\xi+0.2$km/s	-0.00	-0.00	-0.02	-0.01	-0.02	-0.01	-0.02	-0.03	-0.04	0.01	-0.02	-0.01	-0.01	-0.00	-0.00	-0.01
$\xi-0.2$km/s	0.00	0.00	0.02	0.02	0.02	0.01	0.02	0.03	0.05	0.01	0.02	0.01	0.01	0.01	0.00	0.02
Dwarf Totals	0.02	0.08	0.10	0.04	0.04	0.05	0.02	0.10	0.09	0.10	0.02	0.10	0.02	0.11	0.02	0.08
Giants																
$T_{\rm eff}+100K$	-0.02	-0.00	-0.09	0.05	0.04	0.07	-0.02	-0.11	0.08	0.12	-0.01	0.13	-0.01	0.15	-0.03	0.08
$T_{\rm eff}-100K$	0.02	0.00	0.11	-0.05	-0.03	-0.06	0.04	0.10	-0.09	-0.13	0.01	-0.13	0.02	-0.15	0.04	-0.09
log g+0.02	0.00	0.00	0.01	-0.00	0.00	0.00	0.00	0.01	-0.00	-0.00	0.01	-0.00	0.01	-0.00	0.01	-0.00
log g-0.02	-0.00	-0.00	-0.01	-0.00	-0.00	0.00	-0.00	-0.01	0.00	0.00	-0.01	0.00	-0.01	0.00	-0.01	0.00
$\xi+0.2$km/s	-0.00	-0.00	-0.02	-0.02	-0.03	-0.03	-0.04	-0.05	-0.08	-0.01	-0.06	-0.03	-0.07	-0.03	-0.01	-0.03
$\xi-0.2$km/s	0.00	0.00	0.02	0.02	0.04	0.04	0.05	0.05	0.08	0.02	0.07	0.03	0.09	0.03	0.02	0.03
Giant Totals	0.02	0.00	0.11	0.09	0.05	0.08	0.06	0.12	0.12	0.13	0.08	0.14	0.09	0.16	0.04	0.09

Parameter	Cr II	Mn I	Fe I	Fe II	Co I	Ni I	Cu I	Zn I	Y II	Zr I	Zr II	Ba II	Ce II	Nd II	Sm II	
Dwarfs																
$T_{\rm eff}+100K$	-0.03	0.08	0.07	-0.04	0.08	0.06	0.06	-0.00	0.00	0.11	0.00	0.02	0.02	0.02	0.02	
$T_{\rm eff}-100K$	0.03	-0.06	-0.07	0.04	-0.08	-0.06	-0.04	0.01	-0.00	-0.12	-0.00	-0.03	-0.02	-0.02	-0.02	
log g+0.02	0.01	-0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.01	-0.00	0.01	0.01	0.01	0.01	0.01	
log g-0.02	-0.01	-0.00	-0.00	-0.01	-0.00	-0.00	-0.00	-0.00	-0.01	0.00	-0.01	-0.01	-0.01	-0.01	-0.01	
$\xi+0.2$km/s	-0.03	-0.02	-0.03	-0.03	-0.01	-0.03	-0.01	-0.03	-0.02	-0.00	-0.00	-0.07	-0.01	-0.01	-0.00	
$\xi-0.2$km/s	0.04	0.02	0.03	0.03	0.01	0.03	0.01	0.03	0.03	0.00	0.01	0.08	0.01	0.01	0.01	
Dwarf Totals	0.05	0.08	0.08	0.05	0.08	0.07	0.06	0.03	0.03	0.12	0.01	0.09	0.02	0.03	0.02	
Giants																
$T_{\rm eff}+100K$	-0.06	0.10	0.08	-0.07	0.09	0.05	0.05	-0.04	0.00	0.09	-0.01	0.03	0.01	0.02	0.02	
$T_{\rm eff}-100K$	0.05	-0.09	-0.07	0.08	-0.08	-0.03	-0.03	0.06	0.01	-0.09	0.02	-0.02	-0.01	-0.01	-0.02	
log g+0.02	0.00	-0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.01	-0.00	0.01	0.01	0.01	0.01	0.01	
log g-0.02	-0.01	0.00	-0.00	-0.01	-0.00	-0.00	-0.00	-0.01	-0.01	0.00	-0.01	-0.01	-0.01	-0.01	-0.01	
$\xi+0.2$km/s	-0.07	-0.04	-0.07	-0.05	-0.04	-0.06	-0.10	-0.11	-0.07	-0.01	-0.02	-0.07	-0.04	-0.03	-0.03	
$\xi-0.2$km/s	0.08	0.04	0.08	0.05	0.05	0.07	0.13	0.13	0.09	0.02	0.03	0.07	0.06	0.04	0.04	
Giant Totals	0.10	0.11	0.11	0.10	0.10	0.08	0.14	0.14	0.09	0.09	0.04	0.08	0.06	0.04	0.05	

and the count of stars in which a given element was measured. While Fe is measured on the scale of $log_{10}(N(Fe)/N(H)) + 12.00$, all other elements are measured relative to Fe, as denoted by [X/Fe].

TABLE 5.4: Elemental Abundances for NGC 752:
Fe I and Fe II abundances are listed on the scale $log_{10}(N(Fe)/N(H)) + 12.00$ where $log_{10}N(H)$ is set at 12.00, as stated in the text. Listed abundance errors are the star-to-star variations (σ_{var}) and the atmospheric errors (σ_{atm}) from Table 5.3. The atmospheric errors for "All Stars" abundances are the mean of the dwarf and giant stars, weighted by population. The "Q." (quality) score is the average of all measured lines for a given element (see section 4.7.2). Quality scores for synthesized measurements are listed as "n/a".

	All Stars					Dwarfs					Giants				
Ion	Ab.	σ_{var}	σ_{atm}	#	Q.	Ab.	σ_{var}	σ_{atm}	#	Q.	Ab.	σ_{var}	σ_{atm}	#	Q.
Fe I	7.49	0.06	0.08	29	8.0	7.49	0.06	0.08	23	8.0	7.50	0.03	0.11	6	7.9
Fe II	7.50	0.06	0.06	29	8.5	7.49	0.06	0.05	23	8.6	7.53	0.05	0.10	6	7.8
[C I/Fe]	-0.12	0.12	0.13	27	n/a	-0.11	0.12	0.02	21	0.5	-0.22	0.08	0.02	6	n/a
[N I/Fe]	0.23	0.07	0.01	7	n/a	0.13	n/a	0.08	1	1.0	0.27	0.06	0.00	6	n/a
[O I/Fe]	-0.08	0.09	0.10	26	8.3	-0.08	0.09	0.10	20	8.0	-0.08	0.04	0.11	6	9.3
[Na I/Fe]	-0.06	0.12	0.07	20	8.5	-0.12	0.08	0.06	14	8.0	0.11	0.06	0.09	6	10.0
[Mg I/Fe]	0.02	0.07	0.04	29	9.2	0.01	0.07	0.04	23	9.0	0.02	0.06	0.05	6	9.8
[Al I/Fe]	0.11	0.11	0.06	28	9.9	0.09	0.10	0.05	22	9.8	0.22	0.10	0.08	6	10.0
[Si I/Fe]	0.07	0.05	0.03	29	7.6	0.06	0.04	0.02	23	7.6	0.14	0.03	0.06	6	7.7
[Ca I/Fe]	0.02	0.04	0.10	29	6.6	0.02	0.04	0.09	23	6.4	0.03	0.04	0.12	6	7.2
[Sc II/Fe]	0.01	0.09	0.03	28	8.8	0.00	0.10	0.02	22	8.9	0.07	0.01	0.08	6	8.0
[Ti I/Fe]	0.10	0.06	0.11	29	8.2	0.11	0.05	0.10	23	8.3	0.02	0.02	0.14	6	8.0
[Ti II/Fe]	0.11	0.11	0.03	27	8.6	0.12	0.12	0.02	21	8.8	0.07	0.04	0.09	6	8.0
[V I/Fe]	0.12	0.09	0.12	20	8.6	0.13	0.10	0.11	14	8.7	0.08	0.03	0.16	6	8.4
[V II/Fe]	0.21	0.11	0.12	17	7.9	0.22	0.11	0.02	12	7.1	0.19	0.09	0.04	6	9.8
[Cr I/Fe]	0.05	0.06	0.08	29	8.2	0.05	0.07	0.08	23	8.2	0.01	0.02	0.09	6	7.9
[Cr II/Fe]	-0.03	0.06	0.09	13	7.9	-0.02	0.07	0.05	7	8.0	-0.04	0.05	0.10	6	7.7
[Mn I/Fe]	-0.02	0.04	0.09	18	7.0	-0.03	0.04	0.08	12	6.7	0.01	0.01	0.11	6	7.8
[Co I/Fe]	0.10	0.07	0.09	29	8.5	0.10	0.08	0.08	23	8.6	0.09	0.02	0.11	6	8.2
[Ni I/Fe]	-0.03	0.04	0.07	29	7.8	-0.03	0.04	0.07	23	7.8	-0.01	0.01	0.08	6	7.9
[Cu I/Fe]	-0.09	0.05	0.06	22	7.7	-0.09	0.05	0.06	22	7.7	—	—	—	0	—
[Zn I/Fe]	-0.15	0.10	0.05	21	8.0	-0.20	0.05	0.03	15	8.0	-0.01	0.07	0.14	6	8.0
[Y II/Fe]	0.02	0.12	0.05	20	6.6	0.01	0.14	0.03	14	6.1	0.04	0.05	0.09	6	7.8
[Zr I/Fe]	0.17	0.15	0.02	12	7.3	0.27	0.12	0.01	6	6.0	0.08	0.10	0.04	6	8.7
[Ba II/Fe]	-0.14	0.06	0.09	18	6.0	-0.17	0.03	0.09	12	6.0	-0.08	0.07	0.08	6	6.0
[Ce II/Fe]	0.08	0.13	0.04	13	9.2	0.07	0.18	0.02	7	9.5	0.10	0.02	0.06	6	8.7
[Nd II/Fe]	0.21	0.10	0.03	13	9.4	0.25	0.11	0.03	7	9.6	0.17	0.03	0.04	6	9.3
[Sm II/Fe]	0.16	0.11	0.04	10	8.8	0.29	0.05	0.02	4	9.2	0.07	0.02	0.05	6	8.5

5.6.1 Iron and Iron-Peak Elements

I measured the Fe abundance, using over 400 absorption features. As discussed in Section 4.7.2, the large number of measurements was pared to a smaller subset to eliminate data which were likely

affected by measurement errors. Using a weighted average of the calculated Fe I/H (75%) and Fe II/H (25%) species in both dwarf and giant members, I derive a cluster (Fe) metallicity of 7.49 ± 0.06. I then use this value for all atmospheric models and isochrones during the abundance and atmospheric parameter determination process.

In addition to Fe I and Fe II, I also measured the four other "Iron Peak" elements, Cr (I and II), Mn I, Co I, Ni I, and two "nearby" elements V I, and Cu I. Composite cluster values are shown in Table 5.4, and individual line measurements in the digital version of Table A.2. As with Fe abundances, I found no significant differences between abundance measures of these elements in the dwarf and giant population. I accounted for the hyperfine structure of V, Mn, and Cu lines by using the MOOG *blends* driver, with the line component parameters taken from the line sources listed in Table A.1. As with Sc I below, I expected the odd-Z element Co to also have hyperfine structure in its line profiles. However, I had no appropriate source to implement these characteristics, and relied upon an assumed similarity between the NGC-752 stars, and the solar reference counterpart to adjust for this structure.

The measured abundances for the five elements (V I, Cr I, Fe I, Co I, and Ni I) commonly measured in this study and in the giant studies of Carrera & Pancino (2011) and Reddy, Giridhar & Lambert (2012), match well. Table 5.5 (from paper 1) provides a summary of the five elements over the three studies.

Element/Ion	This work	Carrera & Pancino	Reddy, Giridhar, & Lambert
[Fe/H]	-0.01 ± 0.06	$+0.08 \pm 0.04$	-0.04 ± 0.03
[V/H]	$+0.08 \pm 0.03$	$+0.09 \pm 0.09$	-0.01 ± 0.05
[Cr/H]	$+0.01 \pm 0.02$	$+0.08 \pm 0.01$	-0.06 ± 0.04
[Co/H]	$+0.09 \pm 0.02$	$+0.09 \pm 0.04$	-0.05 ± 0.02
[Ni/H]	-0.01 ± 0.01	$+0.07 \pm 0.03$	-0.05 ± 0.03

TABLE 5.5: Comparison of Fe-group abundance measurements between Carrera & Pancino (2011) and Reddy, Giridhar & Lambert (2012) and this work.

5.6.2 Carbon, Nitrogen and Oxygen

While C, N, and O are common in stellar atmospheres, measurement is difficult, even with high S/N spectra, due to the lack of low-excitation transitions in optical spectra. My measurements consist of abundances as calculated from both synthesis and from absorption line feature measurements. For later (nucleosynthetic) analysis in Chapter 7, precise measurements of not only the general elemental abundances are needed, but having isotopic ratios is also helpful. In my NGC 752 spectra I was unable differentiate between isotope states, particularly ^{12}C and ^{13}C, so my analysis is based on relative elemental (instead of isotope) measures.

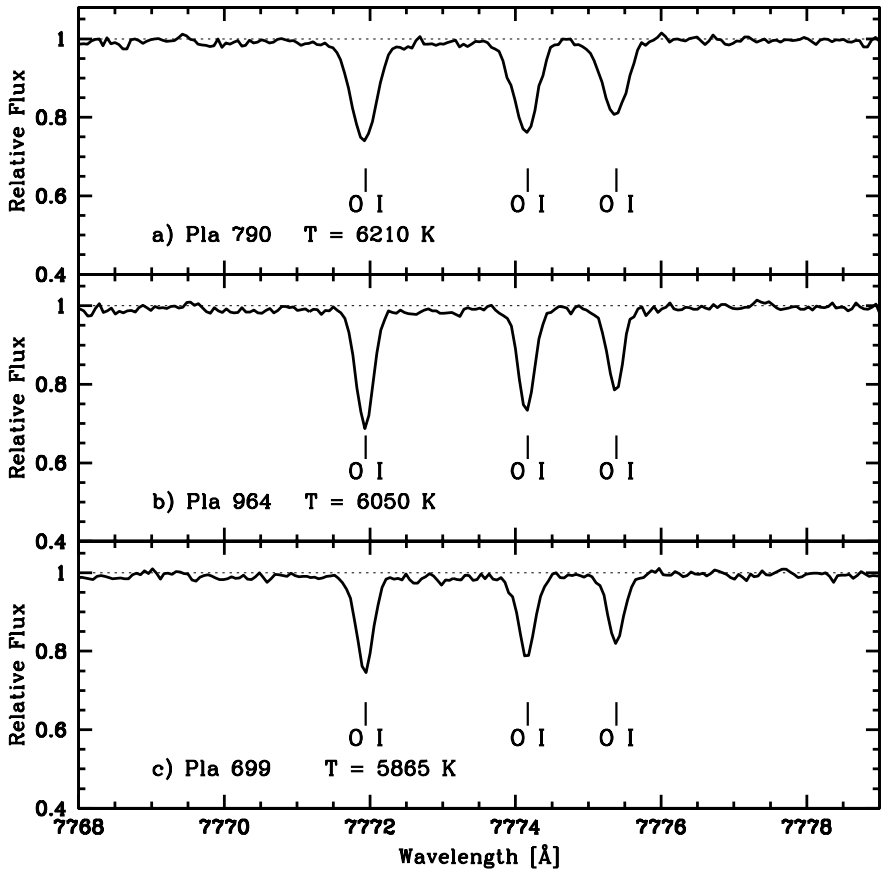


FIGURE 5.4: A sample of three spectra spanning the $T_{\rm eff}$ range for our sample showing the "Oxygen triplet" region (Lum & Boesgaard 2018).

For the giant population, I utilized a combination of elemental and molecular absorption features to calculate C and N abundances. As C-, N-, and O-molecule abundances are inter-related, I first determined the cluster O abundances using equivalent width measurements of several features, including the O I "triplet" at 7771-7776 Å. Figure 5.4 illustrates the spectral region encompassing the triplet region. While MOOG abundance calculations assume Local Thermal Equilibrium (LTE), I adjusted the final abundances for the expected non-LTE (NLTE) conditions of the high-excitation potentials ($\geq 9.14eV$), as suggested in Takeda (2003). I find [O I/Fe]$= -0.08 \pm 0.09$, with identical abundance measurements in the dwarf and giant populations. The cluster, dwarf, and giant (individual star measures in the digital version of the table) [O/Fe] measurements are compiled in Table-5.4. The measured [O/Fe] of -0.08 ± 0.04 in the giant sample is comparable with that of the giants measured by Carrera & Pancino (2011) of $+0.03 \pm 0.04$, who calculated their O abundances using synthesis around the [O I] forbidden line near 6300Å.

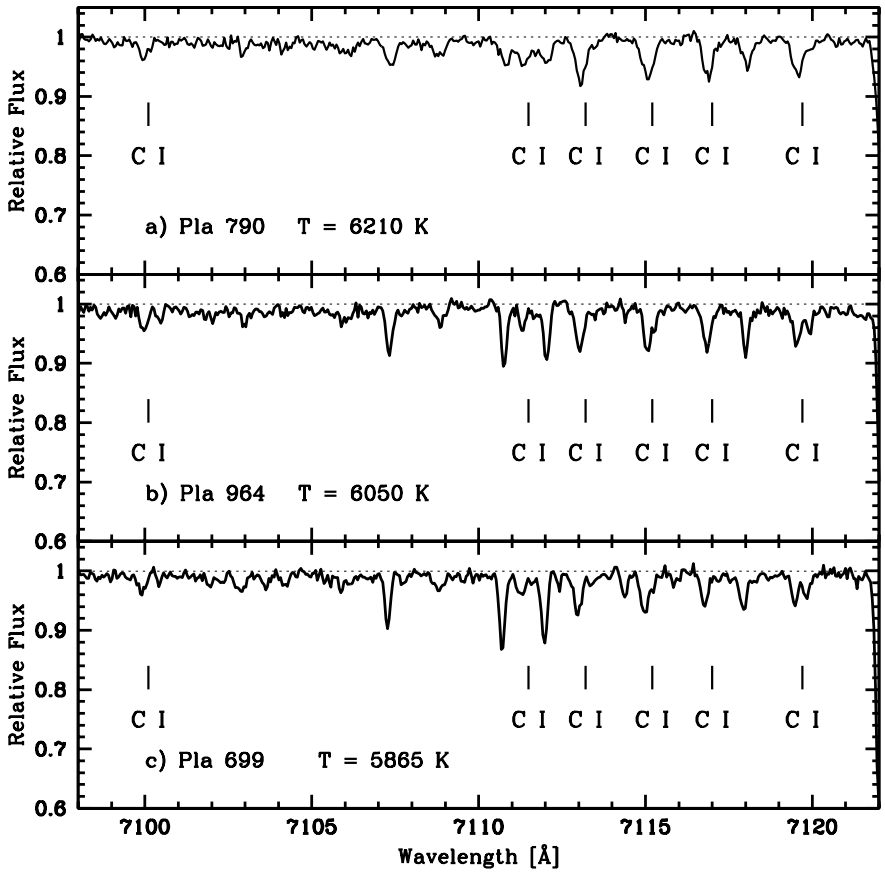


FIGURE 5.5: A sample of three spectra spanning the $T_{\rm eff}$ range for our sample showing the 7110-7120Åregion used for C I synthesis (Lum & Boesgaard 2018).

My abundance determination for C is a composite of equivalent width measurements of the features at 5382 Å, 6381 Å, and 8335 Å and spectral synthesis of the region near 7115 Å, which contains several high-excitation C I absorption features (see Figure 5.5). The synthesis process was partially iterative between C and N abundances, as there are CN features in the same region. In determining the C abundance, I gave higher priority to fitting the lines at 7111 Å and 7113 Å, since they were more isolated from CN features. Table-5.4 lists [C/Fe] results; [C/Fe]$= -0.11 \pm 0.12$ (dwarfs) and [C/Fe]$= -0.22 \pm 0.08$ (giants).

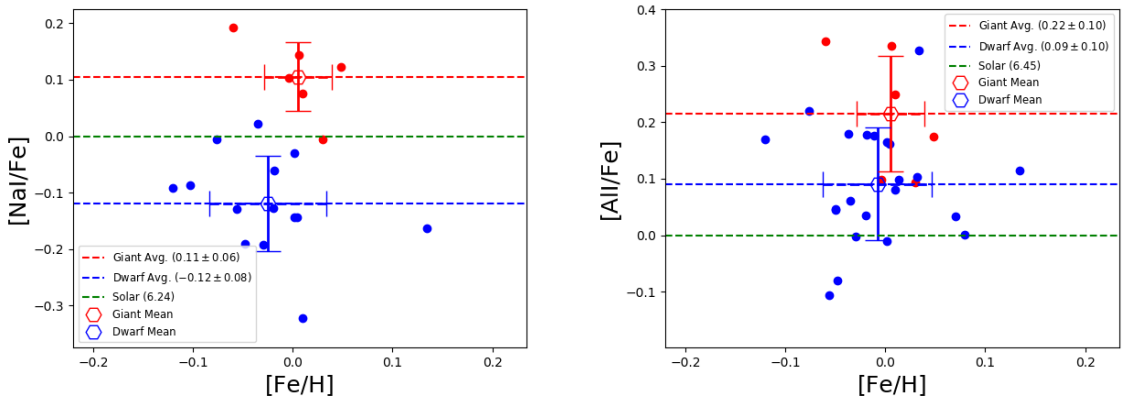
Individual EQW measurements were given equal weight, with the abundance calculated from the synthesis given five (5) times the weight of a single EQW measurement (appropriate for the 5 blended lines in that region). While all of the measured C lines have (relatively) high excitation potentials ($> 7.68eV$), at the effective temperature of our sample stars, I do not expect NLTE effects to be significant (as per Rentzsch-Holm (1996)).

Nitrogen abundance was exclusively calculated using spectral synthesis. As none of the NGC 752 spectra had coverage of the strong CN bandheads in the near-UV and near-IR regions, I used a combination of the N abundance derived during synthesis of the 7115 Å region, and an additional synthesis around the 7442 Å region. Based on the measurement of a single N absorption line in one dwarf spectrum, and synthesis in the five giant spectra, I find [N/Fe]$= 0.13$(dwarfs) and 0.27 ± 0.06(giants)

My measurements of the giant sample show a decrease of C to below that of solar, and N significantly higher. The general pattern of higher N and lower C abundance corresponds to the results of

the giants measured by Böcek Topcu *et al.* (2015), although my measured C abundance is significantly higher, at $[C/Fe] = -0.22 \pm 0.08$ vs. their value of -0.46 ± 0.03. However, our result is nearly identical to the value of $[C/Fe] = -0.23$ from Blanco-Cuaresma *et al.* (2014). The difference is possibly due to the difference in synthesis region selection. Likewise, my measured N abundance of $+0.27 \pm 0.06$ is also higher than that measured by Böcek Topcu *et al.* , which I partially attribute to the unavailability of the violet and red bands of the CN spectrum.

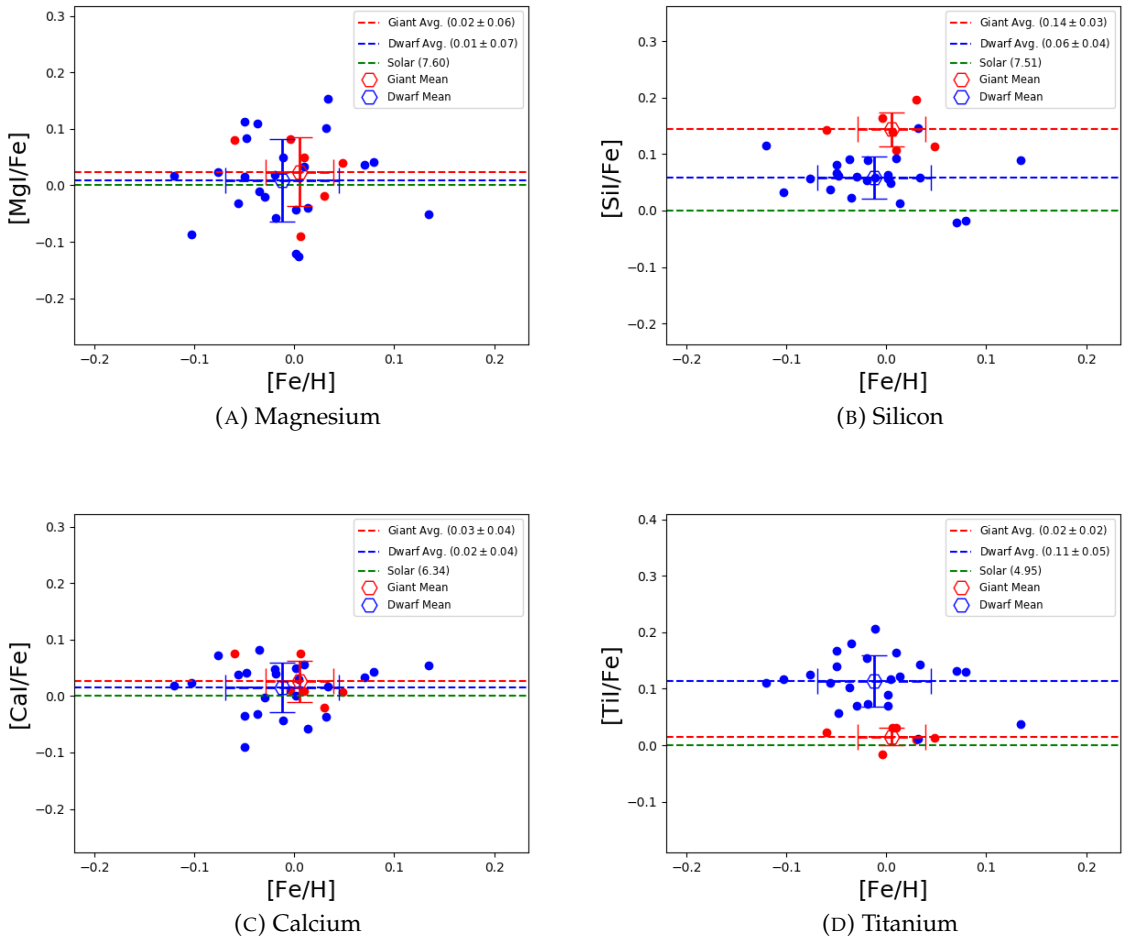
5.6.3 Light and α-Elements



(A) Sodium (B) Aluminium

FIGURE 5.6: Individual star abundances for the odd-Z light elements, Na I, and Al I. The abundance mean for the dwarf population is designated by the blue hexagon, and the giants in red (Lum & Boesgaard 2018).

The "light" elements, with $11 \leq Z \leq 22$, provide viable targets for measurement, even with the lowest S/N spectra. Na, Mg, Al, Si, Ca, and Ti all have either a large number of measurable transitions, or strong, well documented features. In NGC 752, I found that, within error margins, the abundance of these elements follow their solar counterparts. I chose to separate the elements into two groups for discussion. The odd-Z light element sample of Na I and Al I are shown in Figure 5.6. While both light elements show significant star-to-star variation, Na appears to be enhanced in the giants, while Al is likely unchanged. The "α-", or even-Z elements (Mg I, Si I, Ca I, and Ti I), as shown in Figure 5.7, reveal a couple of interesting results. Silicon abundance appears to increase, from 0.06 ± 0.04 in the dwarf population to 0.14 ± 0.03 in the giants, while Ti appears to decrease (0.11 ± 0.05 to 0.02 ± 0.02).



(A) Magnesium (B) Silicon

(C) Calcium (D) Titanium

FIGURE 5.7: Individual star abundances for the alpha-elements, Mg I, Si I, Ca I, and Ti I. The abundance mean for the dwarf population is designated by the blue hexagon, and the giants in red.

While the individual star measurements for Ti I in Figure 5.7d appear to show a clear decrease in Ti abundance in the giant population, I do *not* believe that it accurately represents the physical conditions in the stellar atmospheres. Rather, as noted in a young, pre-main sequence cluster study by D'Orazi & Randich (2009), the discrepancy in Ti I abundance in the cooler giants is a result of over-ionization effects. As in D'Orazi & Randich, we also find that the Ti II abundance in our cooler stars correlate with the same measure in our dwarf population (Table 5.4), and that our giant population with T_{eff} near 5000K, show [Ti II/Fe]-[Ti I/Fe] of approximately 0.1 dex. This is significantly smaller than the ~0.2 dex seen by D'Orazi & Randich, but we note that NGC 752 (~1.5 Gy) is significantly older than IC 2602 and IC 2391 ($<$ 100 My), and we expect our cool giant stars to have less chromospheric activity than the pre-main sequence cool dwarfs in the D'Orazi & Randich study.

5.6.4 Heavy Elements

I have included measurements of several elements with $Z \geq 29$. Generally speaking, high-Z elements track slightly above solar ratios. While s-process enhancement of elements with N=50 (Y) and N=80 (Ba, and Ce to a lesser extent) is possible, the availability of neutron flux relies on higher temperature reactions. These reactions, $^{13}C(\alpha,n)^{16}O$ and $^{22}Ne(\alpha,n)^{25}Mg$, occur at approximately $1.5x10^{8}$ and

$3x10^8K$, respectively (Kaeppeler *et al.* 1990) - temperatures which we do not expect our target star cores to reach. Although, our sample giant stars have not yet reached the Asymptotic Giant Branch (AGB) phase where these reactions could proceed, the He degeneracy fraction prior to the He-flash may allow core temperatures to reach the threshold for these n-generating processes to proceed. However, since we do not measure a significant abundance increase in the end products, O and Mg, of the neutron generating process, we would not expect to see the results from s-processing. Correspondingly, we did not find a significant change in the abundances of the heavy, s-process elements, Y and Ba, and to a lesser extent, Ce, which are included in Table 5.4. Our Sm II abundances, at first glance, would indicate a significant decrease in the abundance of the element. However, this does not hold up to scrutiny - the four values from the dwarf population are based on EQW measurements barely above our measurement threshold of 2mÅ. Similar to the weak N I feature, we did produce five additional measurements of Sm abundances, but did not include them due to the EQW measures falling below our 2mÅ minimum threshold for EQW measurements. If we do include these abundances in our dwarf population, the apparent difference between the dwarf and giant population disappears. Therefore, when measuring Sm in this cluster, we would only rely on EQW measures from the giant population, arriving at a cluster value of [Sm/H]=$+0.07 \pm 0.02$. Individual star abundance measurements on an element-by-element basis are in the digital version of Table 5.4, and the even more detailed individual EQW measures for each line in every star are listed in the included digital material.

Chapter 6

Cross-Sectional Study

The primary data product of this project is a comprehensive elemental catalog of the eight target clusters. Following the process I have described in Chapter 5, this chapter presents a summary of the results from the eight clusters.

6.1 Target Stars

Table 6.1 summarizes the clusters "observed" for this study. The number of spectra which I stacked to produce the science image varied from one to eight, from the instruments discussed in Chapter 4. As with the non-member, and wide-lined binary spectra in NGC 752, I also eliminated a number of spectra to arrive at the final tallies in Table 6.1. Furthermore, since I began the project, the "exclusivity period" has expired and many more spectra for the listed clusters have become available in their respective archives. As an example, there are now spectra for *hundreds* of stars in NGC 2682 (M67) available in the ESO archive, just begging for an extended study of that cluster...

Cluster Name	Age (GY)	Star Counts Total	Giants	Dwarfs	MSTO Mass ($M_\odot$)
Praesepe	0.7	49	3	46	2.3
Hyades	0.8	27	4	23	2.3
IC 4756	1.1	14	13	1	2.0
NGC 752	1.6	29	6	23	1.6
NGC 3680	2.1	10	6	4	1.6
IC 4651	2.5	21	11	10	1.5
NGC 2682 (M67)	3.5	47	12	35	1.3
NGC 6791	8.0	23	10	13	1.1

TABLE 6.1: The list of clusters observed in this work. Ages from Table 3.3

A detailed list of all spectra used for this project (800+) is included as Table A.4. The initial line for each spectrum lists the statistics (S/N, etc.) for the "stack." Individual members of the stack are listed on the subsequent lines.

6.2 Stellar Parameters

I determined atmospheric parameters for all of the target stars using the same iterative process I used to calculate the atmospheric parameters of the NGC 752 targets. Atmospheric parameters for the (200+) target stars are listed in Table A.4, included as a digital table.

6.3 Abundance Catalogs

I used the same automated EQW measuring process as with NGC 752 to measure absorption features in the reduced spectra. A full listing of the measured EQW values as well as the measured line centers is another digital table, Table A.2.

Using the calculated atmospheric parameters, and the EQW measurements, I calculated elemental abundances for all elements in all stars. The resulting table (5000+ lines) is yet another digital inclusion, Table A.5.

6.4 Abundance Analysis

This would be a really boring report if I were to leave the elemental abundance catalog as a simple listing, with no discussion. On the other hand, a star-by-star comparison is needlessly tedious, and statistically insignificant. For the purposes of this study, I have divided my stellar sample by cluster, and then sub-divided by stellar type (dwarf vs. giant). As with the elemental discussion of NGC 752, I have chosen to group elements by a rough "family" categorization, and discuss each in the following sections.

6.4.1 Fe-peak Elements

Cluster	(Stars)	Fe I Ab±σ_{var}	Fe II Ab±σ_{var}	Cr/Fe Ab±σ_{var}	Mn/Fe Ab±σ_{var}	Co/Fe Ab±σ_{var}	Ni/Fe Ab±σ_{var}
Praesepe	(All)	7.61±0.03	7.60±0.04	0.08±0.11	-0.02±0.13	0.48±0.20	0.06±0.09
	(Dwarfs)	7.61±0.03	7.60±0.04	0.05±0.13	-0.01±0.15	0.51±0.21	0.06±0.08
	(Giants)	7.60±0.04	7.43±0.06	0.04±0.06	-0.01±0.07	0.22±0.11	-0.04±0.04
Hyades	(All)	7.61±0.02	7.60±0.07	0.04±0.08	-0.02±0.14	0.08±0.10	0.01±0.07
	(Dwarfs)	7.61±0.02	7.61±0.07	0.04±0.07	-0.04±0.13	0.04±0.11	-0.01±0.05
	(Giants)	7.64±0.01	7.53±0.03	0.06±0.14	-0.01±0.18	0.10±0.18	0.10±0.06
IC 4756	(All)	7.36±0.03	7.33±0.06	0.01±0.11	-0.14±0.16	0.11±0.08	0.06±0.04
	(Dwarfs)	7.33±−	7.30±−	0.02±−	-0.19±−	0.11±−	0.06±−
	(Giants)	7.36±0.03	7.33±0.06	0.02±0.06	-0.13±0.17	0.11±0.08	0.06±0.04
NGC 752	(All)	7.49±0.06	7.48±0.06	0.03±0.06	-0.02±0.04	0.10±0.07	-0.03±0.03
	(Dwarfs)	7.49±0.06	7.47±0.00	0.01±0.02	-0.03±0.04	0.10±0.08	-0.03±0.04
	(Giants)	7.50±0.03	7.51±0.05	0.00±0.05	0.01±0.01	0.09±0.02	0.00±0.01
NGC 3680	(All)	7.45±0.01	7.45±0.06	0.00±0.07	-0.18±0.17	0.31±0.05	0.11±0.08
	(Dwarfs)	7.45±0.02	7.45±0.05	-0.03±0.12	-0.23±0.02	0.33±0.01	0.02±0.02
	(Giants)	7.45±0.01	7.45±0.06	0.00±0.11	-0.16±0.20	0.30±0.06	0.14±0.07
IC 4651	(All)	7.49±0.08	7.49±0.08	0.01±0.12	-0.08±0.18	0.23±0.14	0.06±0.07
	(Dwarfs)	7.50±0.09	7.49±0.05	-0.04±0.14	0.10±0.19	0.24±0.15	0.02±0.06
	(Giants)	7.48±0.05	7.49±0.10	0.06±0.10	-0.19±0.10	0.22±0.09	0.09±0.06
NGC 2682	(All)	7.48±0.03	7.50±0.09	0.00±0.09	-0.09±0.19	0.19±0.15	0.07±0.07
	(Dwarfs)	7.48±0.03	7.49±0.09	0.03±0.09	-0.01±0.15	0.16±0.15	0.06±0.08
	(Giants)	7.48±0.03	7.52±0.09	-0.08±0.08	-0.34±0.38	0.28±0.12	0.09±0.04
NGC 6791	(All)	7.89±0.12	7.86±0.14	0.05±0.14	-0.08±0.18	0.10±0.17	0.06±0.12
	(Dwarfs)	7.87±0.14	7.85±0.22	0.06±0.09	-0.12±0.16	0.16±0.14	0.05±0.09
	(Giants)	7.92±0.12	7.88±0.16	0.04±0.12	-0.05±0.12	0.05±0.12	0.07±0.08

TABLE 6.2: Fe-"peak" elemental abundance measures across the sample of eight clusters, separated by dwarf, giant, and composite cluster abundance.

Table 6.2, above, can be said to have "no surprises." I selected the clusters, based on their (near-) solar metallicity (as measured by Fe abundance), and as would be expected, the Ab(Fe)[1] measurements track close to the Asplund *et al.* solar value of 7.50. The measurements for all of the Hyades and Praesepe Fe-peak elements are comparable (within errors), lending evidence to the "same formation GMC' school of thought. Although, the Co abundances may offer evidence to the contrary, were the uncertainties smaller. NGC-6791 shows Ab(Fe) of 7.88, or [Fe/H] of +0.38, which is again as expected. Recall that NGC 6791 was selected for its age, with the knowledge that it had a higher metallicity than the rest of the study clusters.

For elements with two ionization states (Fe I/Fe II and Cr I/Cr II), when I present an "overall" abundance average for the element, it is calculated as a weighted average, with the unionized species weighted at three times the weight of the singly-ionized species (i.e.:75%/25%).

The remaining Fe-peak elements, Cr, Mn, Co, and Ni all track with their respective cluster Fe abundances. While the mean values of Co (and to a much lesser extent, Cr) seem to show a systematic enhancement over Fe, relative to the solar ratio, the difference is less than the measurement uncertainty.

[1]From Chapter 4: Ab(X) for a given element "X" is a shorthand for $log_{10}(N(X)/N(H)) + 12.00$ where $log_{10}N(H)$ is set at 12.00.

6.4.2 CNO Cycle Elements

Cluster		C I	N I	O I
Praesepe	(All)	0.20±0.31	0.39±0.49	0.14±0.35
	(Dwarfs)	0.23±0.32	0.27±0.43	0.12±0.31
	(Giants)	0.15±0.17	0.65±−	0.19±0.11
Hyades	(All)	0.02±0.18	0.14±0.15	-0.18±0.14
	(Dwarfs)	0.04±0.15	0.08±0.09	-0.12±0.12
	(Giants)	-0.12±0.11	0.45±0.20	-0.28±0.10
IC 4756	(All)	0.11±0.16	0.50±−	0.32±0.13
	(Dwarfs)	0.23±−	—	—
	(Giants)	0.10±0.16	0.50±−	0.32±0.13
NGC 752	(All)	-0.13±0.12	0.25±0.08	-0.08±0.09
	(Dwarfs)	-0.10±0.12	0.13±−	-0.08±0.09
	(Giants)	-0.22±0.08	0.27±0.06	-0.08±0.04
NGC 3680	(All)	0.01±0.18	0.37±0.10	0.11±0.12
	(Dwarfs)	0.03±0.06	—	-0.02±0.02
	(Giants)	-0.01±0.22	0.37±0.10	0.15±0.11
IC 4651	(All)	0.21±0.16	0.27±0.17	0.21±0.13
	(Dwarfs)	0.24±0.17	0.14±0.16	0.24±0.16
	(Giants)	0.17±0.14	0.40±0.09	0.17±0.08
NGC 2682	(All)	0.15±0.22	0.49±0.35	0.17±0.24
	(Dwarfs)	0.14±0.23	0.39±0.19	0.14±0.23
	(Giants)	0.17±0.21	0.54±0.24	0.23±0.25
NGC 6791	(All)	-0.03±0.21	0.23±0.28	-0.08±0.10
	(Dwarfs)	0.04±0.12	0.16±0.06	-0.06±0.08
	(Giants)	-0.09±0.12	0.27±0.23	-0.14±0.14

TABLE 6.3: C, N, and O abundance measures across the sample of eight clusters, separated by dwarf, giant, and composite cluster abundance.

As I discovered with my NGC 752 test case, measurement of C, N, and O, requires high signal-to-noise spectra, optimally including the C line heavy region near 7120Å and the O "triplet" near 7800Å. Unfortunately, neither the HARPS instrument, nor the seemingly "default" UVES instrument setting contain those spectra regions. Nonetheless, I made attempts to measure O lines at 6300Å (and 6363Å in some giants), and C lines near 5380 and 6590Å. Success was...limited. However, when combined with synthesis from spectra where the "good" regions were available, I have sufficient measurements to draw the conclusion that, for all of these clusters, C, and O abundances appear to drop when comparing dwarfs to giants.

Again, as with NGC 752, N abundance measurements are scarce, but also as with 752, N increases from dwarfs to giants. The cause for the C-, and O-depletion, and the N-enhancement is most likely CNO "cycling" as the H->He fusion process in the core during the main sequence lifetime of the star. I'll go into more detail on the evidence for CNO cycling, and nucleosynthesis in the final chapter.

6.4.3 Alpha Elements

Cluster	(Stars)	Ti I Ab±σ_{var}	Ti II Ab±σ_{var}	Mg Ab±σ_{var}	Si Ab±σ_{var}	Ca Ab±σ_{var}
Praesepe	(All)	0.10±0.16	0.15±0.25	-0.12±0.25	0.13±0.12	-0.01±0.14
	(Dwarfs)	0.10±0.16	0.18±0.25	-0.11±0.26	0.13±0.12	-0.01±0.14
	(Giants)	0.06±0.05	-0.06±0.05	-0.29±0.10	0.09±0.05	0.03±0.04
Hyades	(All)	0.13±0.13	0.17±0.13	0.08±0.10	0.11±0.11	0.11±0.07
	(Dwarfs)	0.14±0.13	0.17±0.14	0.07±0.09	0.09±0.06	0.10±0.07
	(Giants)	0.17±0.12	0.20±0.08	0.11±0.12	0.19±0.01	0.12±0.04
IC 4756	(All)	0.04±0.07	0.19±0.10	-0.06±0.09	0.20±0.08	0.10±0.13
	(Dwarfs)	0.08±−	0.13±−	-0.09±−	-0.05±−	0.13±−
	(Giants)	0.03±0.07	0.20±0.10	-0.06±0.09	0.21±0.07	0.10±0.13
NGC 752	(All)	0.10±0.06	0.11±0.11	0.03±0.07	0.07±0.05	0.01±0.04
	(Dwarfs)	0.12±0.05	0.12±0.12	0.03±0.07	0.05±0.04	0.01±0.04
	(Giants)	0.02±0.02	0.07±0.04	0.05±0.06	0.13±0.03	0.02±0.04
NGC 3680	(All)	0.20±0.05	0.27±0.13	-0.06±0.14	0.07±0.10	0.11±0.10
	(Dwarfs)	0.19±0.03	0.21±0.16	-0.12±0.13	-0.04±0.03	0.09±0.02
	(Giants)	0.20±0.05	0.30±0.11	-0.03±0.14	0.11±0.08	0.11±0.11
IC 4651	(All)	0.18±0.19	0.09±0.16	0.05±0.08	0.05±0.08	0.08±0.10
	(Dwarfs)	0.30±0.22	0.04±0.19	0.02±0.09	0.02±0.06	0.07±0.09
	(Giants)	0.08±0.06	0.14±0.14	0.07±0.06	0.09±0.07	0.09±0.11
NGC 2682	(All)	0.24±0.15	0.24±0.15	0.16±0.13	0.13±0.17	0.15±0.13
	(Dwarfs)	0.28±0.14	0.20±0.18	0.17±0.13	0.13±0.19	0.17±0.12
	(Giants)	0.14±0.14	0.36±0.14	0.14±0.12	0.15±0.08	0.07±0.12
NGC 6791	(All)	0.03±0.13	0.05±0.16	0.09±0.15	0.03±0.12	-0.10±0.17
	(Dwarfs)	-0.01±0.08	-0.05±0.12	0.02±0.09	0.07±0.18	-0.14±0.16
	(Giants)	0.05±0.13	0.12±0.14	0.14±0.09	-0.01±0.12	-0.06±0.12

TABLE 6.4: $\alpha-$element (Mg, Si, Ca, Ti) abundance measures across the sample of eight clusters, separated by dwarf, giant, and composite cluster abundance.

My sample of alpha elements, Mg, Si, Ca, and Ti, show very slight enrichment across all eight clusters. While none are quite at the level to be considered "enhanced" (+0.20dex), the fact that all of these clusters show the same slight enhancement, and the fact that they are all fairly old, disk clusters meshes well with the current model of galactic element distribution [***ed: need to find the reference which plots alpha distribution as a fn of radial distance***].

6.4.4 Other Light (Z<22) Elements

Cluster	(Stars)	Na Ab±σ_{var}	Al Ab±σ_{var}	Sc II Ab±σ_{var}	V I Ab±σ_{var}
Praesepe	(All)	-0.04±0.47	0.36±0.35	0.18±0.24	0.32±0.23
	(Dwarfs)	-0.06±0.24	0.38±0.37	0.18±0.25	0.33±0.24
	(Giants)	0.26±0.03	0.18±0.03	0.19±0.08	0.28±0.09
Hyades	(All)	0.02±0.15	0.17±0.12	0.03±0.15	0.19±0.14
	(Dwarfs)	-0.01±0.13	0.17±0.10	0.00±0.14	0.17pm0.14
	(Giants)	0.30±0.05	0.18±0.28	0.12±0.11	0.23±0.18
IC 4756	(All)	0.14±0.13	0.12±0.17	0.18±0.08	0.01±0.07
	(Dwarfs)	-0.12±−	0.20±−	0.09±−	0.10±−
	(Giants)	0.17±0.11	0.11±0.18	0.19±0.08	0.00±0.06
NGC 752	(All)	-0.17±0.14	0.19±0.12	0.01±0.07	0.11±0.09
	(Dwarfs)	-0.25±0.10	0.17±0.11	0.00±0.07	0.12±0.10
	(Giants)	0.01±0.06	0.30±0.10	0.08±0.02	0.08±0.03
NGC 3680	(All)	0.06±0.16	0.20±0.14	0.18±0.13	0.23±0.10
	(Dwarfs)	-0.10±0.10	-0.04±0.01	0.04±0.03	0.33±0.08
	(Giants)	0.11±0.14	0.28±0.04	0.22±0.12	0.20±0.09
IC 4651	(All)	0.13±0.19	0.19±0.17	0.04±0.12	0.23±0.20
	(Dwarfs)	-0.05±0.07	0.04±0.14	0.00±0.10	0.32±0.12
	(Giants)	0.29±0.12	0.30±0.09	0.08±0.11	0.15±0.13
NGC 2682	(All)	0.06±0.19	0.24±0.19	0.13±0.18	0.22±0.21
	(Dwarfs)	0.05±0.13	0.23±0.16	0.10±0.09	0.29±0.13
	(Giants)	0.12±0.16	0.26±0.11	0.21±0.08	0.03±0.17
NGC 6791	(All)	0.19±0.20	0.16±0.18	0.10±0.16	0.05±0.09
	(Dwarfs)	0.10±0.12	0.14±0.16	0.23±0.19	0.08±0.06
	(Giants)	0.28±0.08	0.18±0.16	0.04±0.08	0.02±0.08

TABLE 6.5: The list of clusters observed in this work

As we noted in paper I, and as has been noted in some other works, Na appears to undergo enhancement during a star's main sequence lifetime. Almost all of my spectra contained at least one Na feature, allowing me to gain a fairly good description of this element's evolution. In all cases, I measure an enhancement of at least 0.18dex from dwarfs to giants. Although the difference is slight for my selected lines, I have adopted the Non-Local Thermodynamic Equilibrium (nLTE) adjustments for Na from (Lind *et al.* 2011). For the stars in my study, the main sequence adjustment is on the order of -0.07 to -0.10dex, while the adjustment for giants is slightly larger at -0.10 to -0.12dex.

6.4.5 Heavy (Z>30) Elements

Cluster		Y II	Ba II	Cu	Zn	Zr II	Ce II	Sm II
	(Stars)	Ab$\pm\sigma_{var}$	Ab$\pm\sigma_{var}$	Ab$\pm\sigma_{var}$	Ab$\pm\sigma_{var}$	Ab$\pm\sigma_{var}$	Ab$\pm\sigma_{var}$	Ab$\pm\sigma_{var}$
Praesepe	(All)	0.18±0.23	-0.43±0.31	-0.09±0.38	-0.15±0.41	0.49±0.45	0.45±0.38	0.41±0.36
	(Dwarfs)	0.18±0.22	-0.39±0.29	-0.06±0.39	-0.13±0.41	0.50±0.40	0.50±0.40	0.57±0.37
	(Giants)	0.19±0.06	-0.48±0.29	-0.50±0.43	-0.17±0.31	0.46±0.24	0.38±0.13	0.33±0.24
Hyades	(All)	0.28±0.26	-0.05±0.09	-0.08±0.21	-0.19±0.21	0.21±0.12	0.31±0.36	0.24±0.21
	(Dwarfs)	0.25±0.17	-0.06±0.07	-0.08±0.19	-0.17±0.20	0.21±0.12	0.32±0.37	0.24±0.21
	(Giants)	0.34±0.06	0.17±−	0.11±0.10	-0.21±0.21	—	0.11±0.10	—
IC 4756	(All)	0.24±0.18	0.05±0.15	-0.18±0.18	-0.26±0.15	-0.02±0.26	0.27±0.16	0.02±0.12
	(Dwarfs)	0.26±−	0.14±−	-0.29±−	—	—	-0.10±−	—
	(Giants)	0.23±0.19	0.04±0.16	-0.17±0.38	-0.26±0.25	-0.05±0.24	0.30±0.13	0.02±0.12
NGC 752	(All)	0.03±0.12	-0.03±0.06	-0.07±0.05	-0.14±0.10	0.17±0.15	0.08±0.13	0.16±0.11
	(Dwarfs)	0.02±0.14	-0.10±0.06	-0.07±0.05	-0.20±0.05	0.27±0.12	0.07±0.18	0.29±0.05
	(Giants)	0.06±0.05	-0.01±0.04	—	-0.01±0.07	0.08±0.10	0.10±0.13	0.08±0.02
NGC 3680	(All)	0.52±0.25	-0.23±0.06	0.26±0.16	-0.07±0.26	-0.02±0.15	0.15±0.15	0.01±0.13
	(Dwarfs)	0.28±0.18	-0.23±0.06	-0.02±0.24	—	-0.07±−	—	—
	(Giants)	0.60±0.49	—	0.44±0.22	-0.07±0.26	-0.01±0.33	0.15±0.15	0.01±0.13
IC 4651	(All)	0.16±0.11	0.12±0.12	-0.16±0.32	-0.28±0.14	0.13±0.17	0.02±0.13	-0.10±0.21
	(Dwarfs)	0.13±0.10	0.06±0.11	-0.05±0.31	-0.28±0.12	0.15±0.18	0.21±−	—
	(Giants)	0.18±0.12	0.15±0.11	-0.22±0.33	-0.28±0.23	0.10±0.15	0.00±0.12	-0.10±0.11
NGC 2682	(All)	0.14±0.15	-0.11±0.11	-0.30±0.27	-0.27±0.26	-0.08±0.19	0.03±0.29	0.03±0.30
	(Dwarfs)	0.13±0.16	-0.10±0.16	-0.27±0.35	-0.01±0.12	0.09±0.24	0.06±0.32	0.21±0.33
	(Giants)	0.15±0.20	-0.39±0.34	-0.29±0.42	-0.15±0.25	-0.21±0.29	-0.03±0.18	-0.09±0.21
NGC 6791	(All)	0.10±0.12	0.05±0.10	—	—	—	—	0.06±0.10
	(Dwarfs)	0.08±0.09	0.06±0.08	—	—	—	—	—
	(Giants)	0.16±0.10	0.01±0.09	—	—	—	—	0.06±0.10

TABLE 6.6: "Heavy" element ($Z > 29$) abundance measures across the sample of eight clusters, separated by dwarf, giant, and composite cluster abundance.

Table 6.6 lists cluster-by-cluster abundance calculation results for the sample set. The combination of the low relative abundance of these elements in a stellar atmosphere, combined with only a few well-understood (mainly uncertainty in the "Logg_f" physical parameter) absorption features for each element makes a precise automated measurement difficult. However, a cursory analysis does reveal a couple of interesting points

First, while the Y II "excess" abundance over the solar ratio was not significant in the case study of NGC 752, analysis of the stars in all eight of the sample clusters does indicate that there is a general trend of excess Y II in local clusters - or, more likely, that our Sun is (very slightly) deficient in Y II, in comparison to the local cluster population. A similar, but less conclusive, over-abundance (or under-abundance in the solar atmosphere) exists in the Ba II measurements. I discuss the potential for Y II and Ba II enhancement due to s-process nucleosynthesis in the next chapter.

One of the measurements which I feel best demonstrates the potential of my automated process is the Sm II measurements. Sm II is extremely rare in the solar atmosphere (Ab(Sm)= 0.96, Asplund *et al.*), but I was able to detect, Sm II features in the giant atmospheres on a fairly regular basis. While I

don't use the calculations in the s-process analysis in the next chapter, Sm II abundance changes could be used as a comparison against Y II and Ba II for s-process analysis, in the future.

Chapter 7

Conclusions

Over the past few chapters, I have presented a number of new processes, and analysis using these processes. These are new techniques for optimizing stellar analysis, adding easily an order of magnitude to the number of samples a single astronomer can analyze. As the available data archives continue to add observations, the amount of spectra available for cluster (and other stellar) studies will grow exponentially. Just as the days of manually tracing spectra disappeared in favor of digital analysis, so will the practice of manually measuring spectra also go by the wayside. The software packages included with this dissertation places me at the forefront of this new field.

This final chapter is a sort of "executive summary" of the highlights of my research, detailed in the prior chapters. I have also added a summary of the evidence for nucleosynthetic processes in my observed clusters.

7.1 Automated Analysis

Chapter 4 introduced a software package which can take a high-resolution spectral product from a telescope archive, and produce a comprehensive elemental abundance catalog for the target star's atmosphere.

Due to my choice of a central database, open source licensing, Python code, and central source repository and code control, future additions to the code capabilities will be easy. For instance, should a future researcher wish to measure an element or ionization state not currently covered, they would simply need to enter the new absorption line physical parameters into the SQLite database, and re-run the code on their target spectra. No actual code changes are necessary.

A similar process could be use to explore more star population abundance scenarios. As an example, while I measured about 50 stars from NGC 2682 (M67), since I began the project, the ESO has released spectra for literally hundreds of stars in that cluster. As I stated in my introduction, cluster abundance research has been limited to typical sample sizes of a handful to a few tens of stars in a cluster. To produce such analysis on hundreds of stars in a cluster is unprecedented...and is definitely something I would put into the category of "Future Work."

7.2 Machine Algorithm for Cluster Membership

Chapter 3 presented a new method for determining cluster membership, which takes advantage of the excellent photometry of the GAIA spacecraft. This method uses advanced machine learning techniques to analyze position and motion characteristics and find cluster members.

While I had the most enjoyment designing and executing this part of my research, I had to ask myself: "Is this method any better, or different than the commonly used methods?" I found the answer when reading Gaia Collaboration *et al.* (2018), which created cluster membership maps using the "traditional" methods of hard position and motion limits. I found that my cluster membership lists were on the order of twice the size of the GAIA groups'. The majority of my "extra" members lay outside of the positional limits of the GAIA group. My external members also lie outside of the sphere of gravitational influence of the cluster. In the narrowest terms, this would eliminate them as cluster

members. However, recall that as a cluster ages, (lower mass) stars escape the gravitational bounds of the system. Many of these stars will not have velocity significantly above that required to escape the cluster. They will remain near the cluster, in terms of galactic position, moving away at a velocity not much different than the cluster mean velocity.

My probabilistic model is adept at finding these "moving group" members. Thus, the "extra" members in my cluster membership determinations. Unfortunately, it was outside of the scope of my project to pursue this line of research, but it is another field I would like to add to the "Future Research" pile.

7.3 Elemental Evolution

After producing the comprehensive elemental catalog, the secondary goal of this project was to look for evidence of alteration in stellar composition due to fusion processes during the main-sequence lifetime of the stars. Of the known processes, I found strong observational evidence of two, and possible evidence for a third.

7.3.1 Carbon, Nitrogen and Oxygen

With the assumed age range between 700 Myr and 8 Gyr, and Main Sequence Turn Off (MSTO) masses between $2.4M_{\odot}$ and $1.2M_{\odot}$, I expected the evolved members of my cluster stars to show evidence of CNO-cycle processing. Even though the CNO processing is ongoing throughout a star's lifetime, conditions at the core-radiative zone boundary isolate core materials until core hydrogen fusion ends. Convective processes during the H shell fusion, and core He fusion phases allow the altered core materials to mix into the stars' atmosphere, becoming detectable in giant star spectra. Iben (1964) and Iben (1991) summarize the atmospheric differences I expected to see as a result of this "first dredge-up". Specifically, there should be increases in ^{13}C (relative to ^{12}C) and ^{14}N.

Böcek Topcu *et al.* explained the relative N enhancement and C depletion to main-sequence evolutionary processes, namely He production in the core through the CNO cycle. The relevant portion of the CNO cycle is that which occurs at the lowest temperature:

$^{12}C(p,\gamma)^{13}N(,e^{+}\gamma_{e})^{13}C(p,\gamma)^{14}N(p,\gamma)^{15}O(,e^{+}\gamma_{e})^{15}N(p,\alpha)^{12}C$

The slow step, or "bottleneck", is the ^{14}N proton capture. Over the course of the main-sequence lifetime of a star, this results in an increase of N abundance at the expense of the ^{12}C abundance. With the IC 4756, Hyades, and Praesepe MSTO masses above $2.0M_{\odot}$, evidence of processing from the higher temperature branch of the CNO-I-cycle (CNO-II) might also be present. Specifically, the high temperature branch has two steps which require proton capture by O. Like the N increase in the CNO-I cycle, increased O abundance might be explained through the capture steps in the higher temperature CNO-II cycle. I did not measure a significantly increased O abundance over any of the clusters, including the three with the highest turnoff masses. Therefore, I conclude that these higher temperature branches are not significant contributors to the elemental alteration of the sample clusters.

Admittedly, measuring the CNO abundances in giant stars alone only shows half the picture. The critical comparison should occur between main-sequence and evolved members. To that end, I measured C and O abundances using the absorption features at 7115 Å and 7775 Å, respectively. These abundance measurements do hint at the expected trend for CNO processing, but measurement uncertainty, particularly due to the lack of strong CN features in our dwarf sample, prevents me from making a definitive statement to that effect, based on C and O measures, alone.

The third "piece" - N abundance - shows stronger evidence of the expected trend for CNO processing. N abundance measurements are mostly the result of molecular CN feature synthesis, so I was largely unable to determine N abundances for a large fraction of the dwarf population. The few dwarf N abundance measurements usually are a single EQW measurement of the 7442.29 Å N I feature. Frequently, the software was able to detect the 7442.29 Å line, but disqualified it due to the EQW measurements falling below the acceptable minimum width threshold for the spectra's S/N ratio. However, in the few cases where I have dwarf and giant N abundance measures, they show a marked increase from dwarfs to giants.

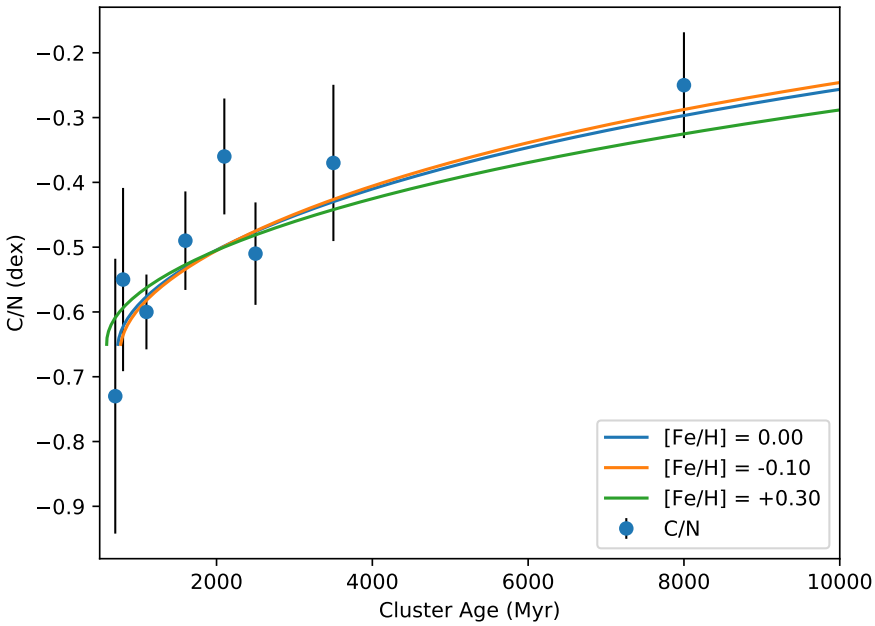


FIGURE 7.1: C/N ratios in the giants of my selected clusters (plotted by age). Included in the figure are the analytical fits for metallicity ([Fe/H]) values of -0.10, 0.00, and +0.30 from Salaris *et al.* (2015).

In paper I we noted that the [C/N] ratio of -0.49, as measured in our NGC 752 giant sample, agreed with the expected [C/N] of -0.54 for a 1.5Gyr giant just after its first dredge-up ($[C/N]_{FDU}$) at solar metallicity given in Salaris *et al.* (2015). Salaris *et al.* provide an analytical solution for the Age-(C/N, [Fe/H]) relation, which I have plotted in Figure 7.1 along with all of my target cluster [C/N] ratios. Although my error bars are large, and the Salaris relation is not well defined below 900Myr, my measurements do appear to follow the expected trend.

7.3.2 Na, Mg, Al, and Si - Evidence for nucleosynthesis

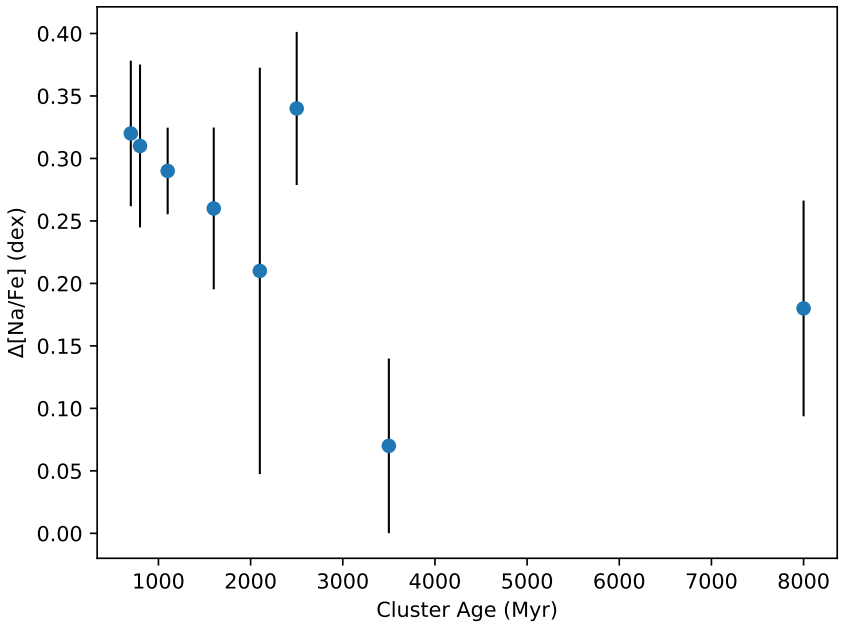


FIGURE 7.2: The increase of Na abundance when comparing giant to dwarf atmospheres, defined as Δ[Na/Fe], plotted against cluster age. The points show definite Na engancement in the giants, decreasing with cluster age, possibly due to lower core temperatures in less massive stars.

When comparing Na abundances between the dwarf and giant populations, I found that Na abundance shows a clear increase between the dwarf and giant populations, as shown in Figure 7.2, above. For simplicity, $\Delta Na = A(Na_{\rm giant}) - A(Na_{\rm dwarf})$. While there does appear to be a trend in the mean ΔNa, when uncertainties are taken into account, it is hard to make a claim on a temporal trend. Even though my choice of Na lines did not include the strong 5890/5896 Å or 8183/8195 Å doublets, I did correct for NLTE effects, as suggested in Lind *et al.* (2011). The measured Na increase in the giants is possibly due to the same mechanism as Boesgaard, Roper & Lum (2013) noted in comparing their Praesepe dwarf population with the giants studied by Carrera & Pancino (2011). Specifically, that the enrichment is a result of the NeNa cycle as detailed in Arnould & Mowlavi (1995). As with the O and N enhancement during CNO processing, the ^{23}Na proton capture has the smallest cross-section, which provides an explanation for my measured enhancement. The relevant cycle is:

$$^{20}Ne(p,\gamma)^{21}Na(e+\gamma_e)^{21}Ne(p,\gamma)^{22}Na(e+\gamma_e)^{22}Ne(p,\gamma)^{23}Na(p,\alpha)^{20}Ne$$

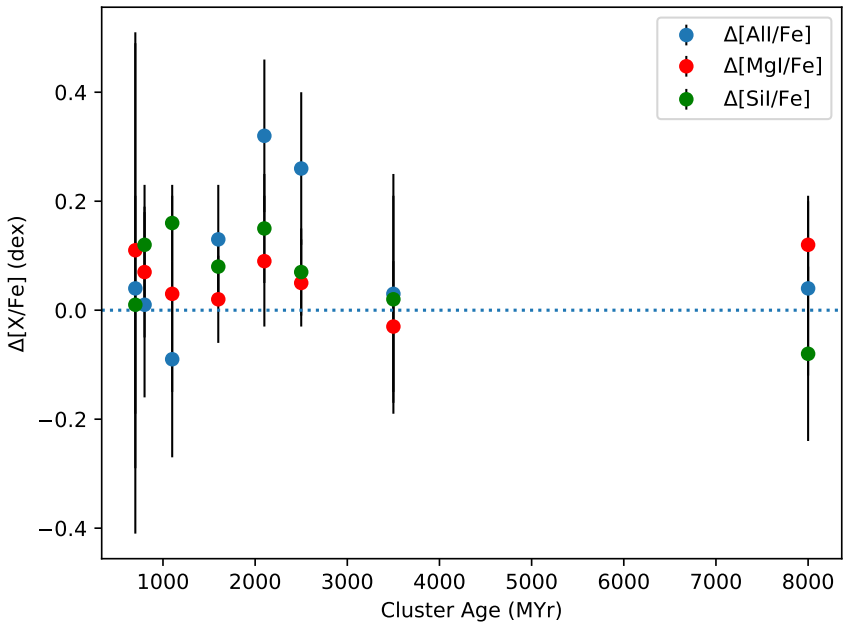


FIGURE 7.3: Mg, Al, and Si abundances changes (defined as Δ[X/Fe]), plotted against cluster age. While the data might indicate an enhancement of these light elements, the trend is not as clear as in Figure 7.2, above.

Arnould & Mowlavi and Ventura, Carini & D'Antona (2011) also discuss Al enhancement through the MgAl cycle. While Ventura, Carini & D'Antona specifically address the MgAl cycle in AGB stars, the conditions in the core during the "He-flash" can easily reach the necessary conditions for the process to occur. However, when comparing my dwarf and giant populations, the measured Al abundances differ by less than the star-to-star variance, and do not show a consistent pattern between clusters. Similarly, I did not measure significant decrease in Mg, the depleted element in the MgAl cycle, between my two populations. As with Boesgaard, Roper & Lum, we may even measure an *increase* in Mg in the giants - again only if uncertainties are ignored. This result wouldn't be unexpected, as the $^{23}Na(p,\gamma)^{24}Mg$ step within the NeNa cycle can provide a "leakage" of Mg, resulting in an abundance increase. Ventura, Carini & D'Antona also state that increases in Si as a result of the MgAl cycle are possible. As can be seen in Figure 7.3, none of these three light elements show clear evidence of enhancement over the Main-Sequence life of their clusters (stars).

7.4 Future Work

In sections 7.1 and 7.2, above I discussed the two "obvious" extensions to my work. In addition to those, the other "obvious" extension project would be to increase my sample of clusters. This study was limited to nearby, approximately solar metallicity, "middle-aged" clusters. There's no real reason to limit an extended project by metallicity or age. Distance is only an issue in obtaining spectra which, due to the archive-based "observations" of this research, is only a short-term limit. As CCD and instrument technology advances, and as larger, and next-generation space telescopes come online, the distance limitation will be less of an issue.

In the immediate future, four of my target clusters lay in Kepler observation "fields", NGC 6791, Praesepe, Hyades, and NGC 2682 (M67). Eager planet hunters will, no doubt, take more high-resolution

spectra for radial velocity (variation) measurements - which would in turn provide an opportunity to add more spectra to my existing sample.

Finally, a wise engineer once said "No software project is ever complete." My software has many opportunities for improvement. The prime example can be seen by the fact that my full list of absorption features contains almost 2500 features, while a typical spectra will have less than 1000 usable measured features. My software aggressively (and correctly) prunes what it considers "bad" measurements, due to excessive blending, poor fitting, and weak features. Many of these discarded measurements could be retained with a more accurate, and "smarter" algorithm.

Other potential software features could include an integrated archive search and spectral preparation package. Instead of manually searching archives, a user could simply ask for spectra from a star, and an automated search algorithm could page through known archives. Basically, the potential for software feature additions is endless.

Appendix A

Digital Table Samples

A large amount of data was used in analyzing my stellar spectra. For the sake of space, I have included the largest data in digital (VOTable) format. I include samples of the tables here.

GAIA Source #	Mem	Literature ID	Mem	RA	Dec	pm_{RA} (mas/yr)	pm_{Dec} (mas/yr)	Parallax	Membership Reference	Cluster
050298125783225088	62	19082	100	4 : 05 : 25.8	19 : 26 : 31.2	115.1	-32.7	21.14 ± 0.05	Perryman *et al.* (1998)	Hyades
047019347749289216	48	19098	100	4 : 05 : 39.8	17 : 56 : 15.3	122.1	-27.4	21.83 ± 0.04	Perryman *et al.* (1998)	Hyades
045367056650753280	66	19148	100	4 : 06 : 16.3	15 : 41 : 52.9	114.5	-19.8	21.03 ± 0.04	Perryman *et al.* (1998)	Hyades
				...						

TABLE A.1: Example entries from the cluster membership table. Included as a digital data supplement in Table A.1

Ion	Wavel. (Å)	Ex. Pot. (eV)	$log(gf)$	Ref.
C I	5380.340	7.68	-1.62	58
C I	6587.620	8.53	-1.00	23
		...		

TABLE A.2: Example entries from the absorption feature physical parameter table ("Line List"). Included as a digital data supplement in Table A.2

Ion	Wavel. (Å)	Spectra Filename	Meas. WL.(Å)	EQW (mÅ)
22.1	7564.4894	0752_PLA_0429_KH_000.fits	7564.488	1.984
22.1	6827.9362	0752_PLA_0429_KH_000.fits	6827.968	1.865
13.0	6698.670	Hyades_HIP_13976_HA_001.fits	6698.659	50.63
13.0	6696.032	0752_PLA_0790_KH_000.fits	6696.116	28.56
		...		

TABLE A.3: Example entries from the measured Equivalent Widths (EQW) table. Included as a digital data supplement in Table A.3

Star			Wavel.	S/N		Obs.	
Cluster	ID	Instrument	Range (Å)	6200 Å	7200 Å	Date	P.I
NGC-3680	EGG-013	HARPS	3781-6913	65	—		HARPS
Hyades	HIP-20978	HIRES	3872-6943	210	—	2012-02-27	
			...				

TABLE A.4: Example entries from the full list of spectra used for this research. Included as a digital data supplement in Table A.4

Cluster	ID	T_{eff} (K)	log g	V_{turb} (km s^{-1})
Hyades	HIP-19793	5750	4.45	1.2
NGC-0752	PLA-0828	5475	4.53	1.9
		...		

TABLE A.5: Example entries from the table of stellar parameters used in this project. Included as a digital data supplement in Table A.5

Star		[FeI/H]				[FeII/H]				[C/Fe]				
Cluster	ID	[FeI/H]	σ_{var}	#	Q.	[FeII/H]	σ_{var}	#	Q.	[C/Fe]	σ_{var}	#	Q.	...
NGC-2632	KW-335	7.64	0.18	29	8.0	7.57	0.18	3	8.7	-0.08	0.02	2	6	...
						...								

TABLE A.6: Example entries from the table of all elemental abundances measured for a given star. Included as a digital data supplement in Table A.6

Appendix B

Cluster Membership Evaluations

B.1 Coma

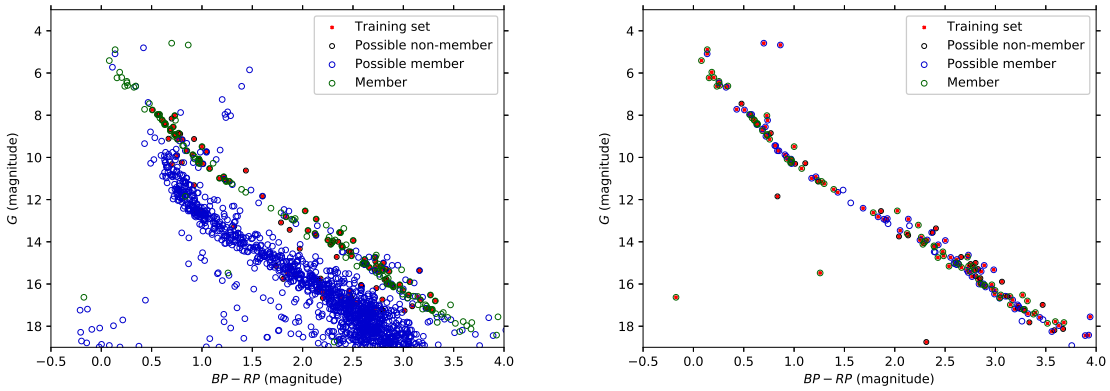


FIGURE B.1: Color-Magnitude Diagram evolution for the Coma data set. The presence of an apparent second, background cluster is discussed in Section 3.3.1. Color and point markers are the same as in Figure 3.3.

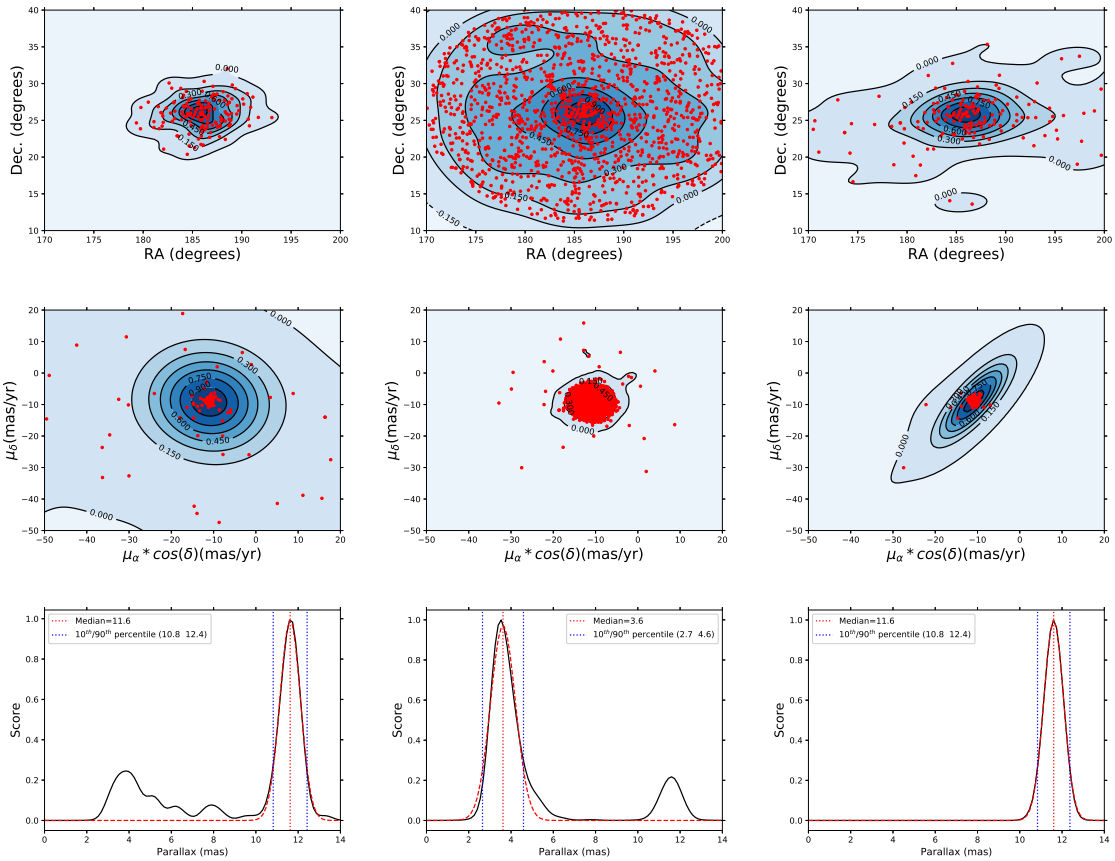


FIGURE B.2: Coma cluster KDE evolution using: The initial (literature) member set (left). Highest scoring member set after the first iteration (center). Highest scoring members after the second and final iteration (right).

B.2 Hyades

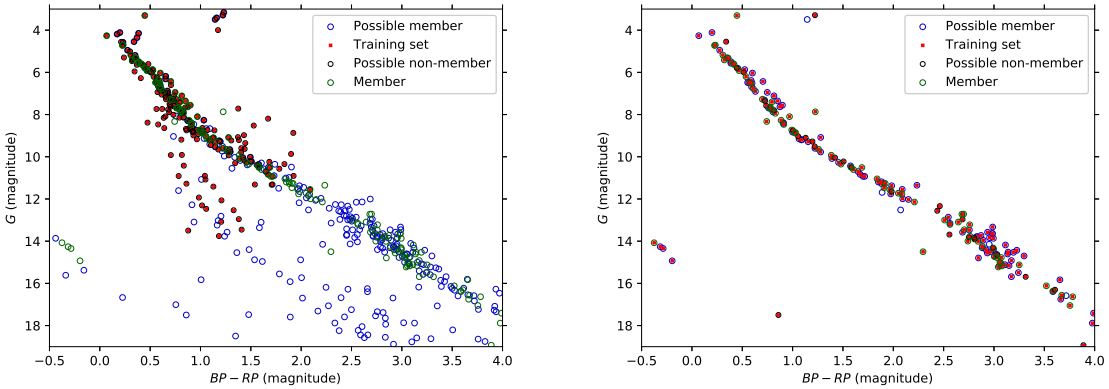


FIGURE B.3: Color-Magnitude Diagram evolution for the Hyades data set. Color and point markers are the same as in Figure 3.3.

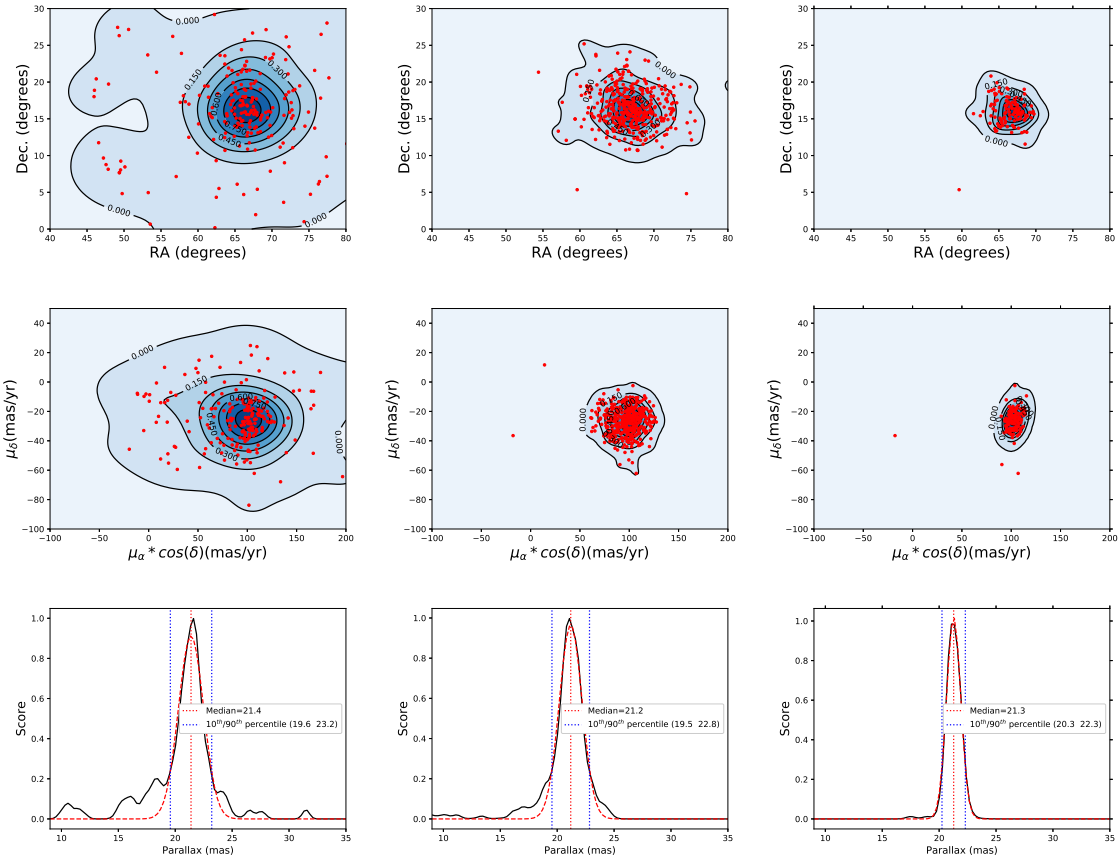


FIGURE B.4: Hyades cluster KDE evolution, using the same population distributions as in Figure B.2

B.3 IC 4651

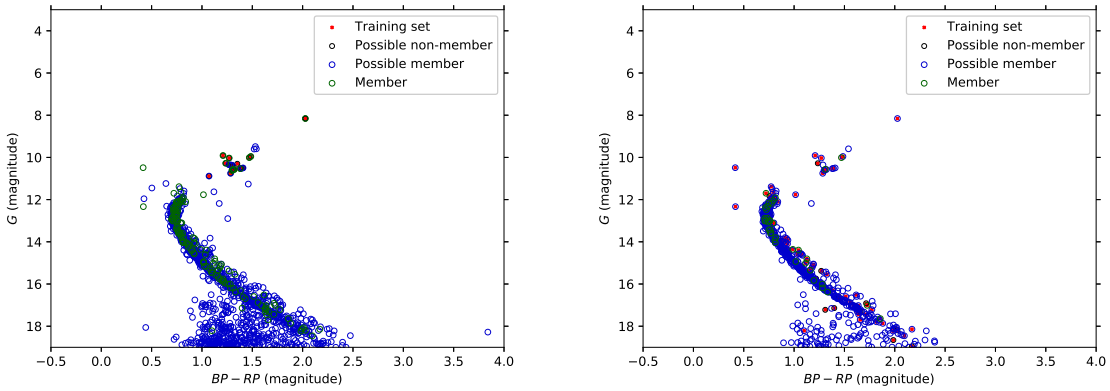


FIGURE B.5: Color-Magnitude Diagram evolution for the IC 4651 data set.. Color and point markers are the same as in Figure 3.3.

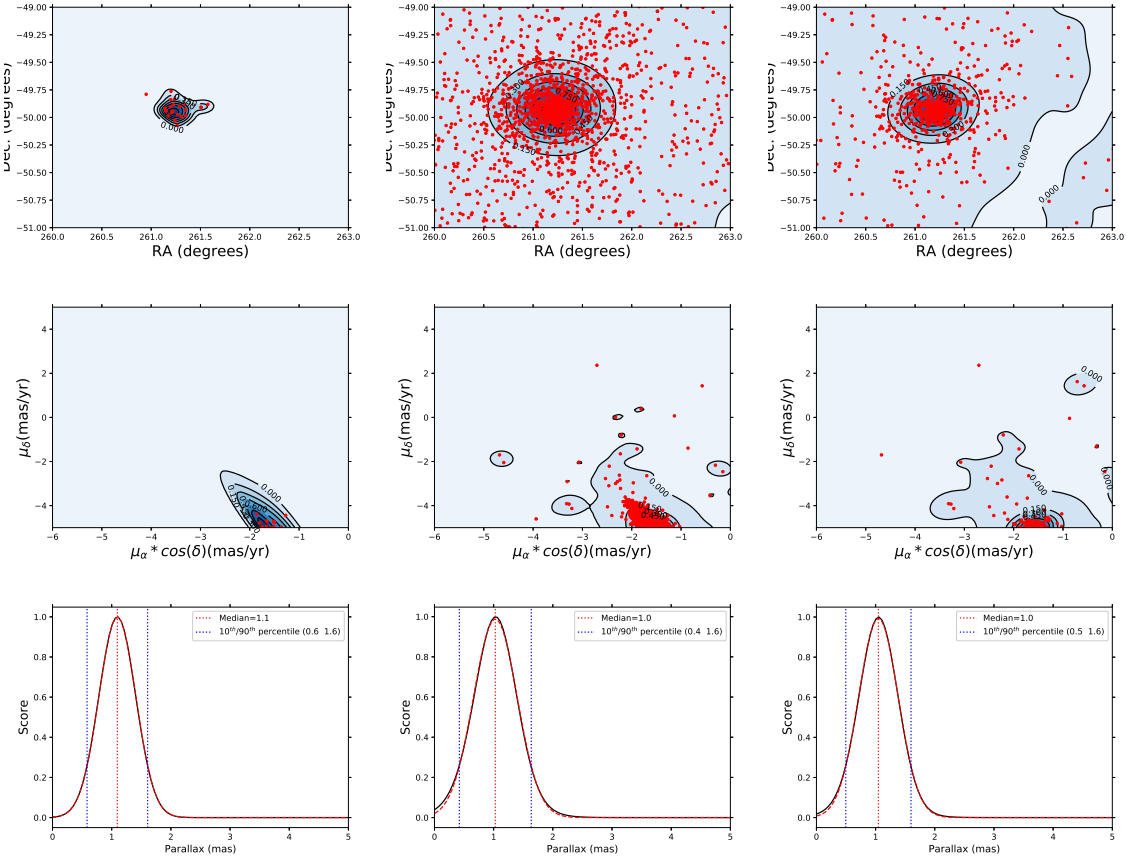


FIGURE B.6: IC 4651 cluster KDE evolution, using the same population distributions as in Figure B.2

B.4 IC 4756

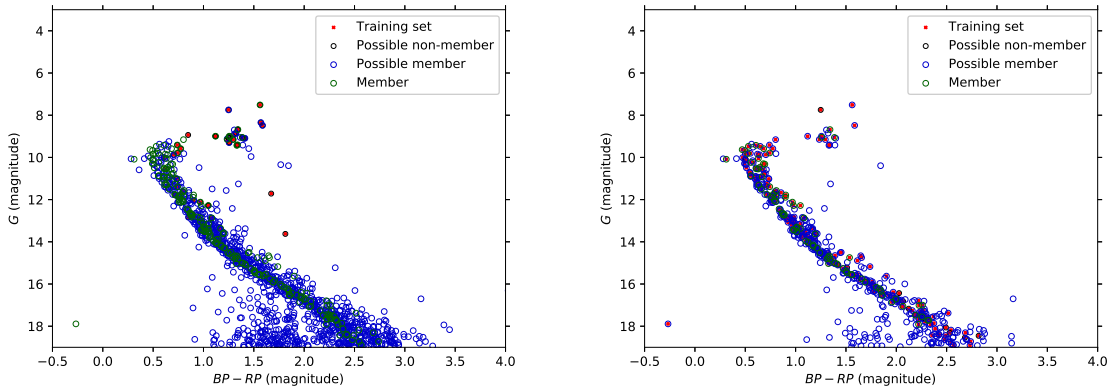


FIGURE B.7: Color-Magnitude Diagram evolution for the IC 4756 data set. Color and point markers are the same as in Figure 3.3.

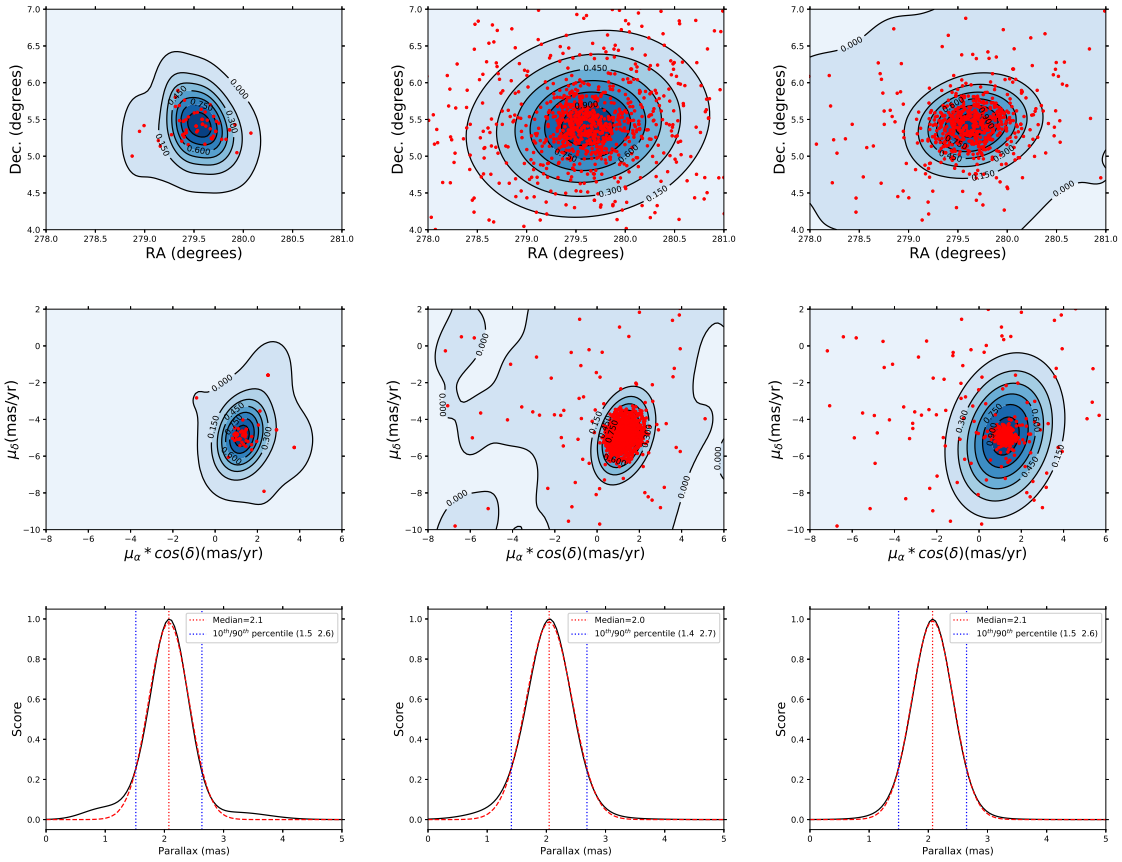


FIGURE B.8: IC 4756 cluster KDE evolution, using the same population distributions as in Figure B.2

B.5 M7

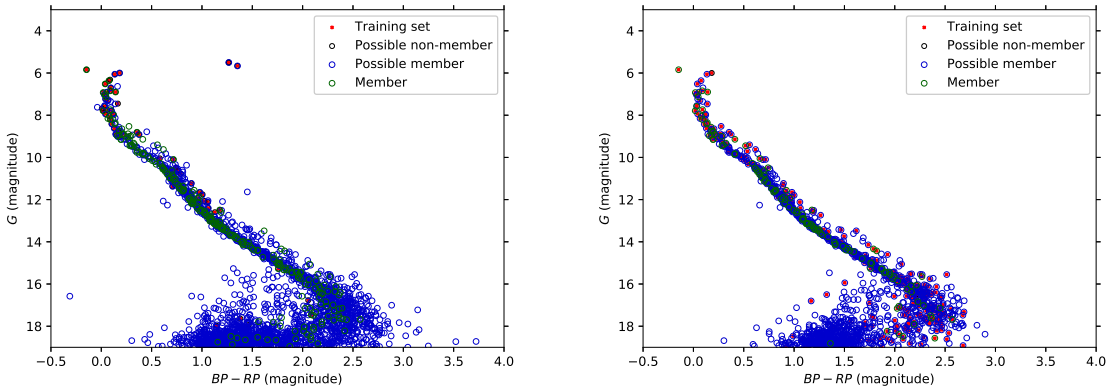


FIGURE B.9: Color-Magnitude Diagram evolution for the M7 data set. Color and point markers are the same as in Figure 3.3.

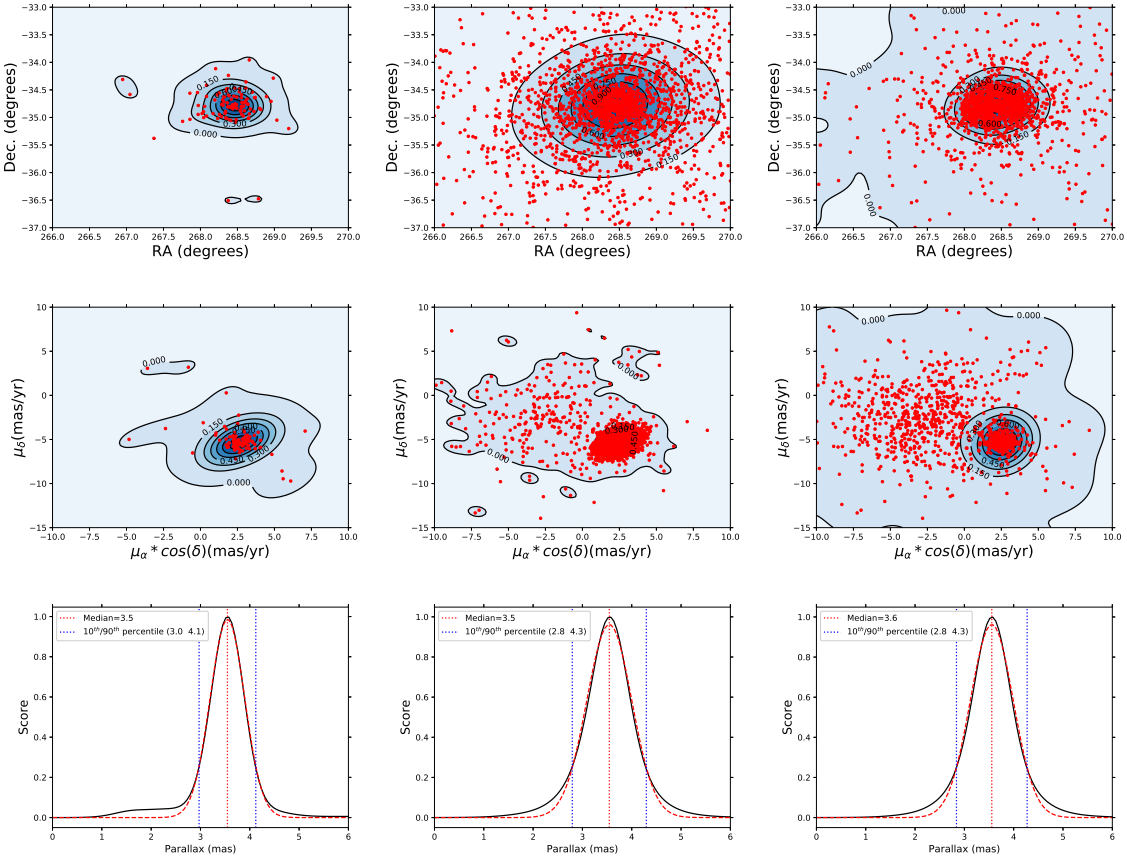


FIGURE B.10: M7 cluster KDE evolution, using the same population distributions as in Figure B.2

B.6 M67

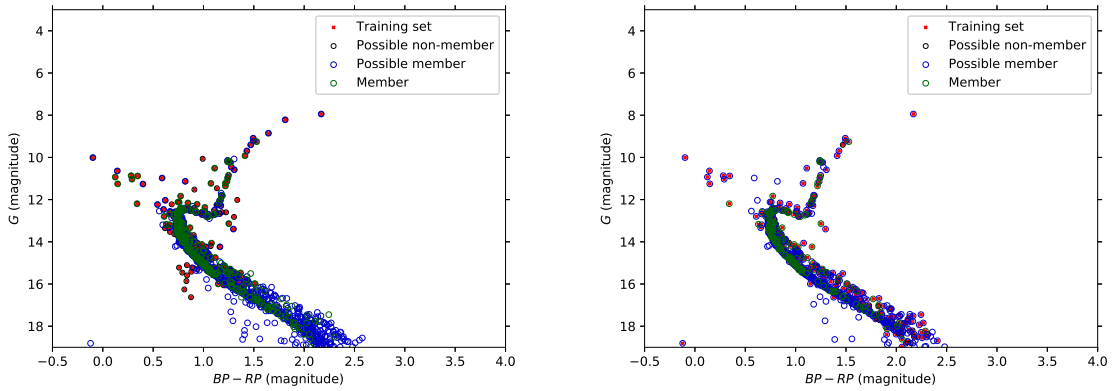


FIGURE B.11: Color-Magnitude Diagram evolution for the M67 data set. Color and point markers are the same as in Figure 3.3.

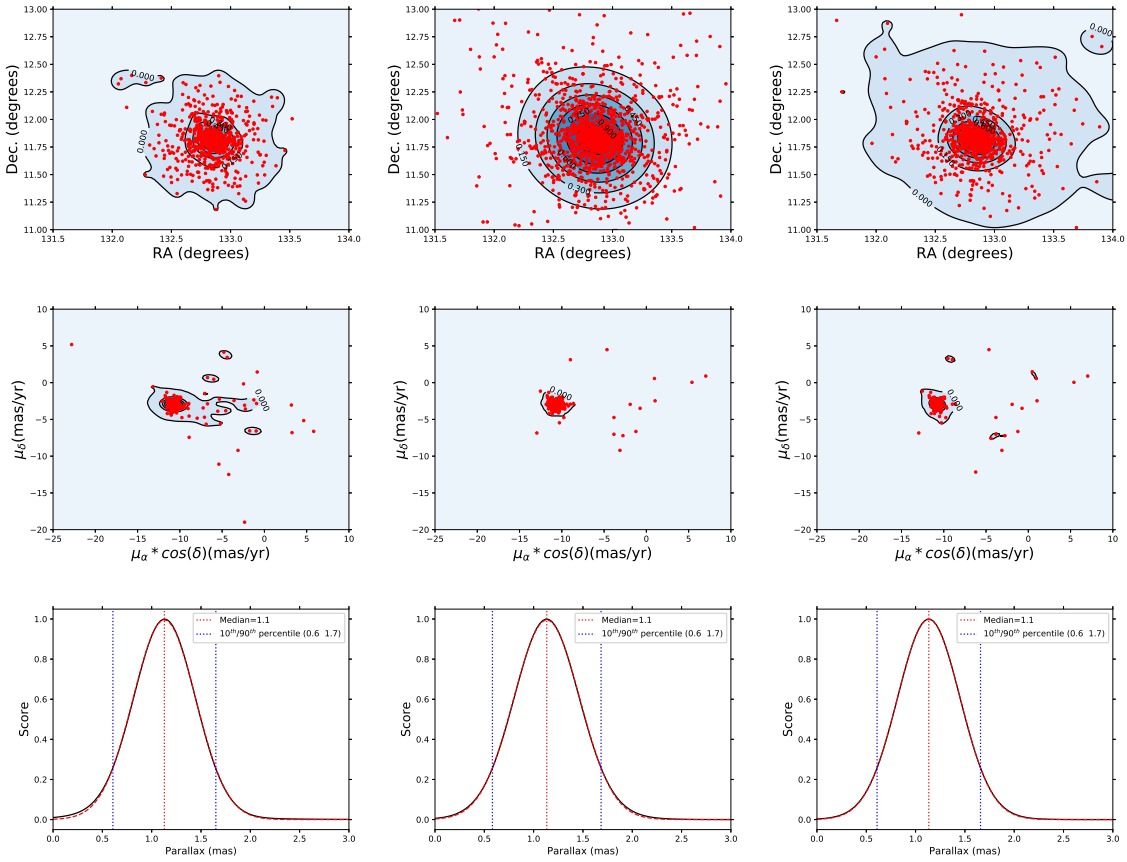


FIGURE B.12: M67 cluster KDE evolution, using the same population distributions as in Figure B.2

B.7 NGC 752

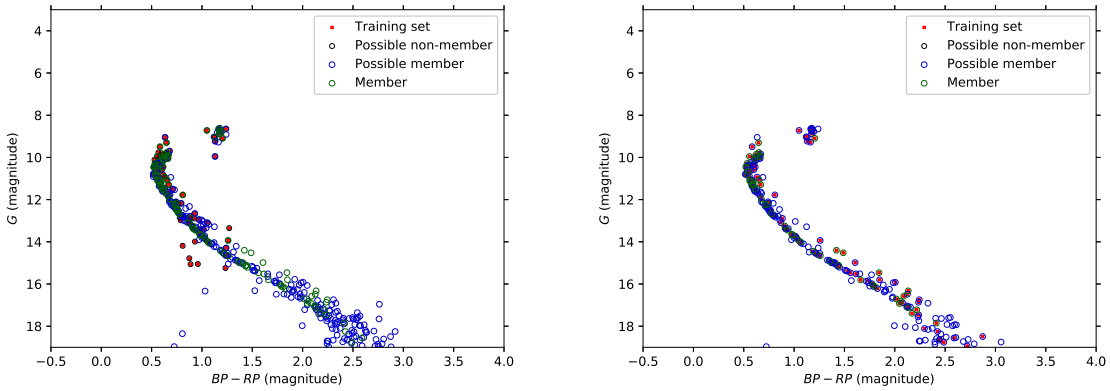


FIGURE B.13: Color-Magnitude Diagram evolution for the NGC752 data set. Color and point markers are the same as in Figure 3.3.

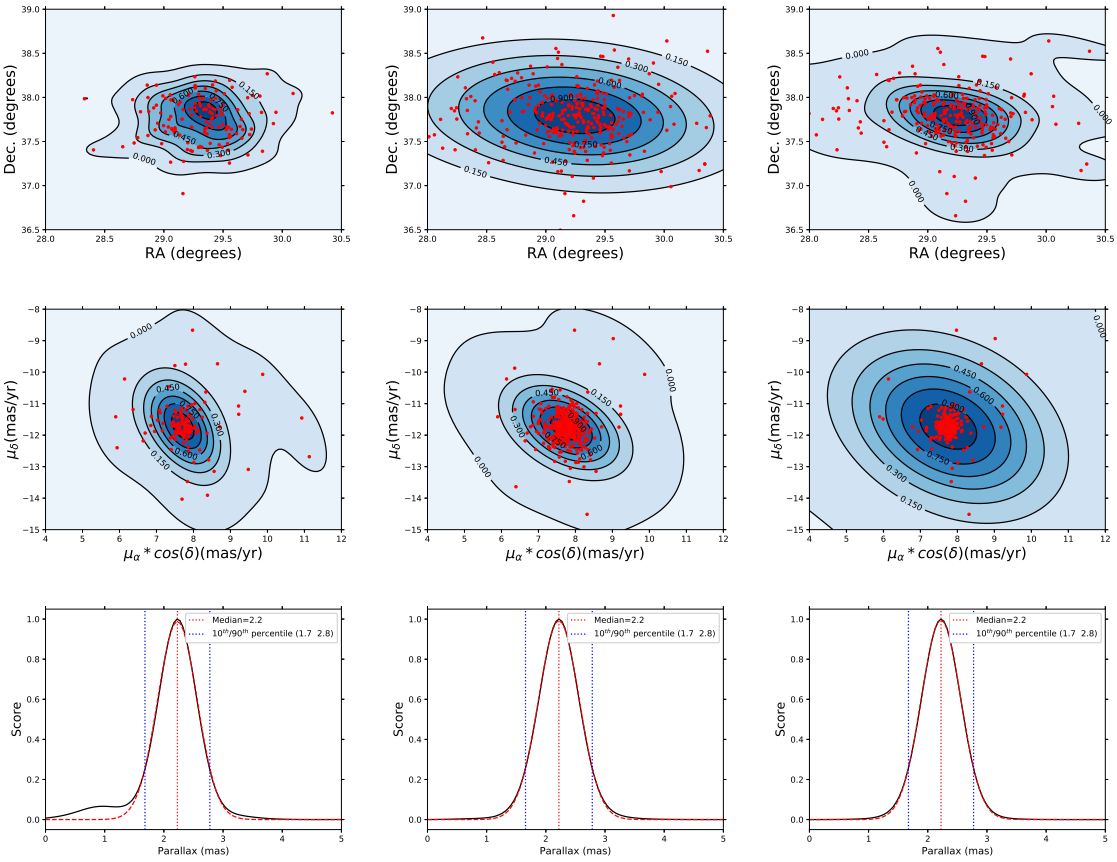


FIGURE B.14: NGC 752 cluster KDE evolution, using the same population distributions as in Figure B.2

B.8 NGC 3680

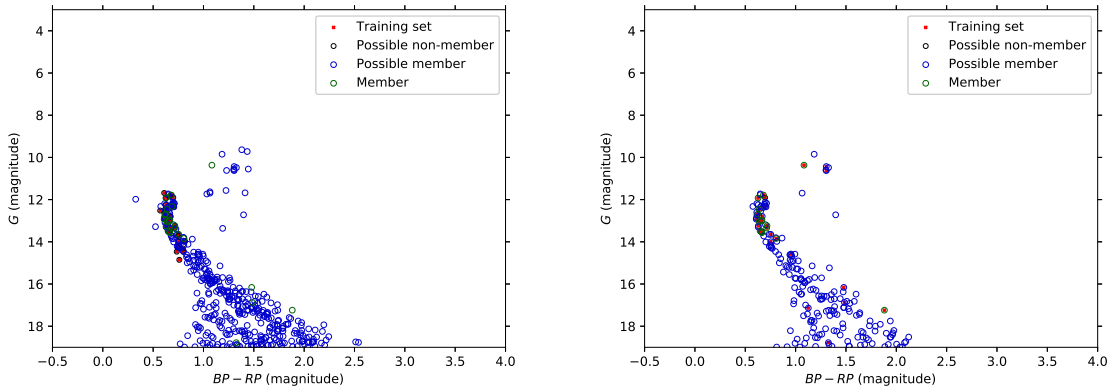


FIGURE B.15: Color-Magnitude Diagram evolution for the NGC3680 data set. Color and point markers are the same as in Figure 3.3.

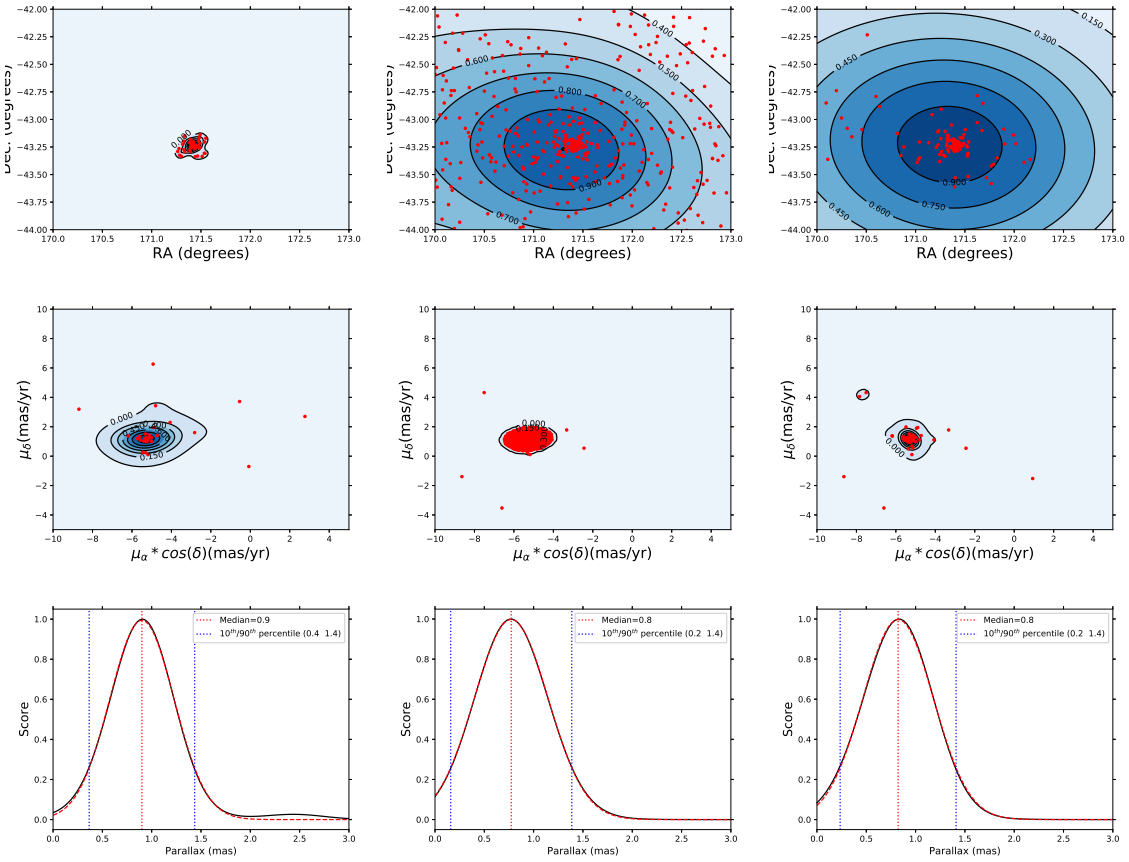


FIGURE B.16: NGC 3680 cluster KDE evolution, using the same population distributions as in Figure B.2

B.9 NGC 6791

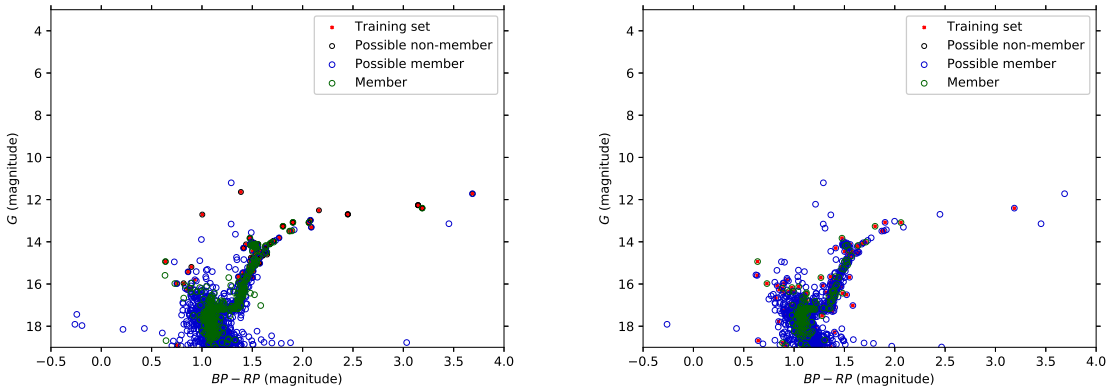


FIGURE B.17: Color-Magnitude Diagram evolution for the NGC 6791 data set. Color and point markers are the same as in Figure 3.3.

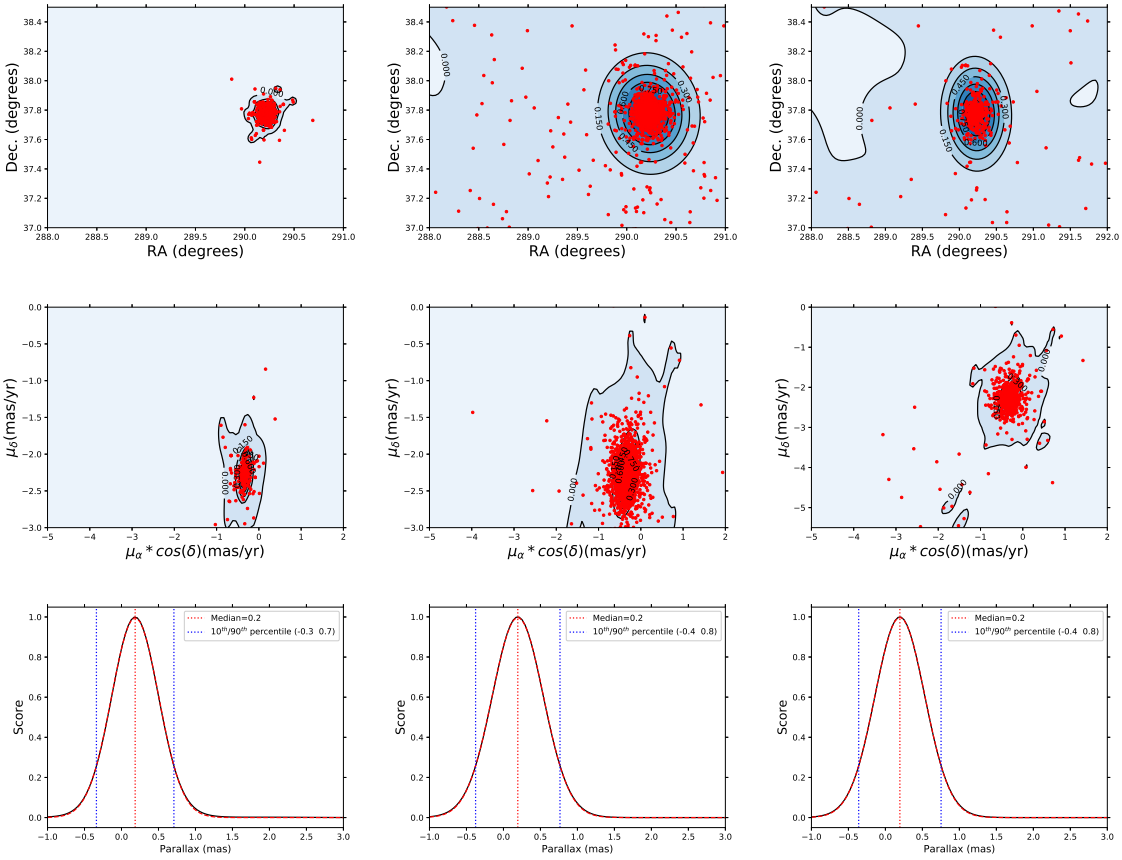


FIGURE B.18: NGC 6791 cluster KDE evolution, using the same population distributions as in Figure B.2

B.10 Praesepe

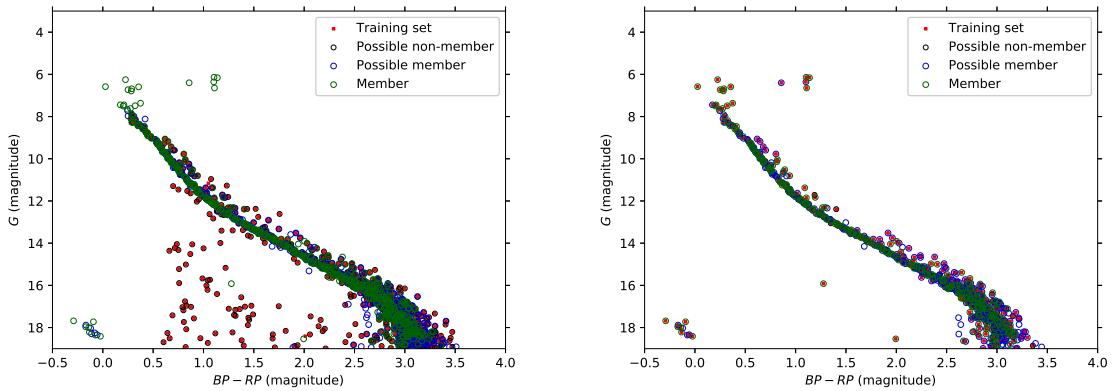


FIGURE B.19: Color-Magnitude Diagram evolution for the Praesepe data set. Color and point markers are the same as in Figure 3.3.

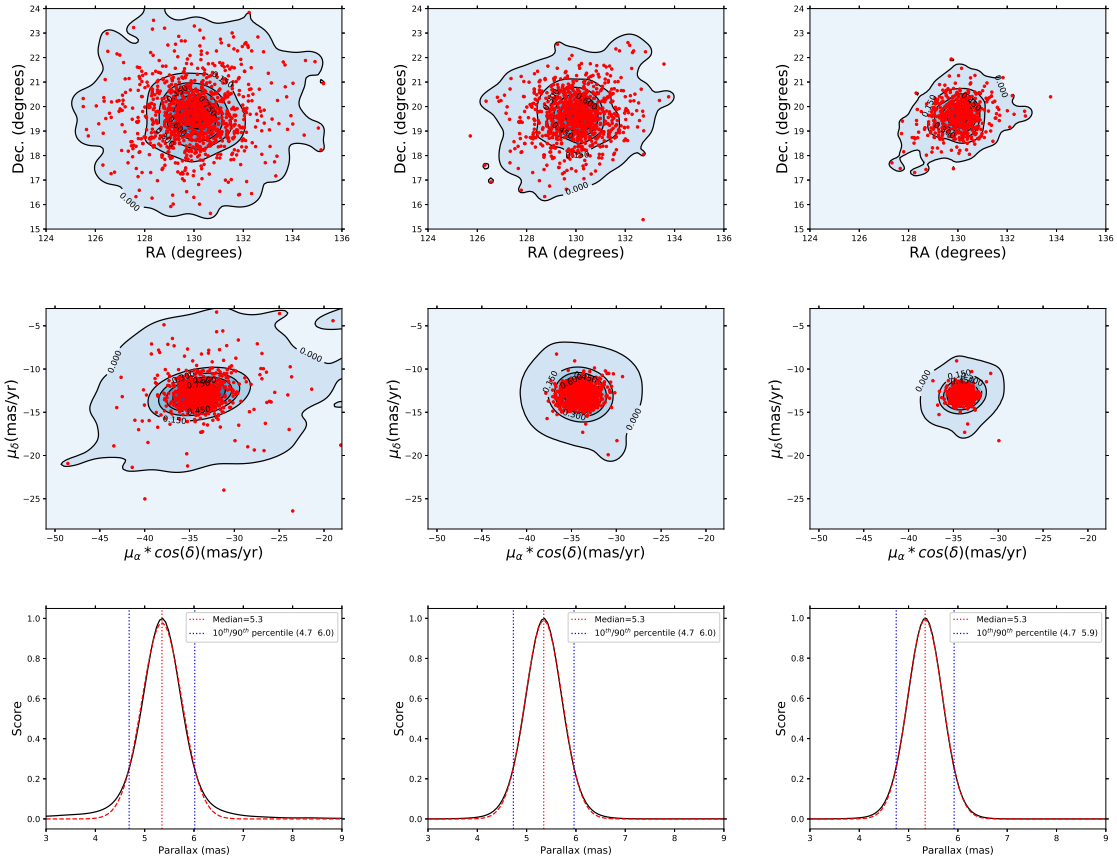


FIGURE B.20: Praesepe cluster KDE evolution, using the same population distributions as in Figure B.2

Appendix C

Absorption Line References

The absorption lines listed in the digital inclustion of Table A.1 come from >60 literature sources. This appendix contains the list of those sources.

TABLE C.1: Absorption Line References

Ref. No.	Reference	Ref. No.	Reference
1	Allen & Porto de Mello (2011)	35	Garz (1973)
2	Allende Prieto & Garcia Lopez (1998)	36	Giridhar & Arellano Ferro (1995)
3	Bard, Kock & Kock (1991)	37	Hannaford *et al.* (1982)
4	Bard & Kock (1994)	38	Kock & Richter (1968)
5	Beveridge & Sneden (1994)	39	Kramida *et al.* (2014)
6	Biemont *et al.* (1991)	40	Kroll & Kock (1987)
7	Biemont & Godefroid (1980)	41	Lambert & Warner (1968)
8	Bizzarri *et al.* (1993)	42	Lambert & Luck (1978)
9	Blackwell, Shallis & Simmons (1980)a	43	Lawler & Dakin (1989)
10	Blackwell *et al.* (1986)a	44	Lawler *et al.* (2013)
11	Blackwell *et al.* (1986)b	45	Lawler *et al.* (2008)
12	Blackwell *et al.* (1986)c	46	Lawler, Sneden & Cowan (2015)
13	Blackwell, Menon & Petford (1983)	47	Lawler *et al.* (2009)
14	Blackwell *et al.* (1982)a	48	Lawler *et al.* (2014)
15	Blackwell, Petford & Simmons (1982)d	49	Martin, Fuhr & Wiese (1988)
16	Blackwell *et al.* (1976)	50	McWilliam & Rich (1994)
17	Blackwell *et al.* (1979)a	51	McWilliam *et al.* (1995)
18	Blackwell, Petford & Shallis (1979)b	52	Meylan *et al.* (1993)
19	Blackwell *et al.* (1980)b	53	Moity (1983)
20	Blackwell *et al.* (1982)b	54	Nissen & Schuster (1997)
21	Blackwell *et al.* (1982)c	55	O'Brian *et al.* (1991)
22	Blackwell, Menon & Petford (1984)	56	Prochaska *et al.* (2000)
23	Boesgaard, Roper & Lum (2013)	57	Prochaska *et al.* (2000)
24	Booth *et al.* (1984)	58	Ramírez & Allende Prieto (2011)
25	Buurman *et al.* (1986)	59	Reddy, Giridhar & Lambert (2013)
26	Cardon *et al.* (1982)	60	Savanov, Huovelin & Tuominen (1990)
27	Cohen *et al.* (2003)	61	Schnabel, Kock & Holweger (1999)
28	Den Hartog *et al.* (2014)	62	Smith & Raggett (1981)
29	Den Hartog *et al.* (2011)	63	Sobeck, Lawler & Sneden (2007)
30	Edvardsson *et al.* (1993)	64	Stephens (1999)
31	Francois (1988)	65	Wickliffe & Lawler (1997)
32	Fry & Carney (1997)	66	Wiese & Martin (1980)
33	Fuhr, Martin & Wiese (1988)	67	Wood *et al.* (2013)
34	Fuhrmann, Axer & Gehren (1995)	68	Wood *et al.* (2014)

Bibliography

Allen, D. M. and Porto de Mello, G. F. 2011, A&AS, 525, A63.

Allende Prieto, C. and Garcia Lopez, R. J. 1998, A&AS, 131, 431.

Anthony-Twarog, B. J. *et al.* 2009, AJ, 138, 1171.

Anthony-Twarog, B. J. *et al.* 1991, AJ, 102, 1056.

Anthony-Twarog, B. J. and Twarog, B. A. 2004, AJ, 127, 1000.

Anthony-Twarog, B. J., Twarog, B. A., and Mayer, L. 2007, AJ, 133, 1585.

Arnould, M. and Mowlavi, N. 1995, in Liege International Astrophysical Colloquia, volume 32 of Liege International Astrophysical Colloquia, 17.

Asplund, M. *et al.* 2009, ARA&A, 47, 481.

Astropy Collaboration *et al.* 2013, A&A, 558, A33.

Bard, A., Kock, A., and Kock, M. 1991, A&A, 248, 315.

Bard, A. and Kock, M. 1994, A&A, 282, 1014.

Bedin, L. R. *et al.* 2008, ApJ, 678, 1279.

Bedin, L. R. *et al.* 2006, A&A, 460, L27.

Beveridge, R. C. and Sneden, C. 1994, AJ, 108, 285.

Biazzo, K. *et al.* 2007, Astronomische Nachrichten, 328, 938.

Bica, E. *et al.* 2001, A&A, 366, 827.

Biemont, E. *et al.* 1991, A&A, 249, 539.

Biemont, E. and Godefroid, M. 1980, A&A, 84, 361.

Bizzarri, A. *et al.* 1993, A&A, 273, 707.

Blackwell, D. E. *et al.* 1986a, MNRAS, 220, 549.

Blackwell, D. E. *et al.* 1986b, MNRAS, 220, 303.

Blackwell, D. E. *et al.* 1986c, MNRAS, 220, 289.

Blackwell, D. E. *et al.* 1979, MNRAS, 186, 633.

Blackwell, D. E. *et al.* 1976, MNRAS, 177, 219.

Blackwell, D. E., Menon, S. L. R., and Petford, A. D. 1983, MNRAS, 204, 883.

Blackwell, D. E., Menon, S. L. R., and Petford, A. D. 1984, MNRAS, 207, 533.

Blackwell, D. E. *et al.* 1982a, MNRAS, 201, 611.

Blackwell, D. E., Petford, A. D., and Shallis, M. J. 1979, MNRAS, 186, 657.

Blackwell, D. E. *et al.* 1982b, MNRAS, 199, 21.

Blackwell, D. E. *et al.* 1980, MNRAS, 191, 445.

Blackwell, D. E. *et al.* 1982c, MNRAS, 199, 43.

Blackwell, D. E., Petford, A. D., and Simmons, G. J. 1982, MNRAS, 201, 595.

Blackwell, D. E., Shallis, M. J., and Simmons, G. J. 1980, A&A, 81, 340.

Blanco-Cuaresma, S. and Soubiran, C. 2016, in SF2A-2016: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics, 333.

Blanco-Cuaresma, S. *et al.* 2015, A&A, 577, A47.

Blanco-Cuaresma, S. *et al.* 2014a, A&A, 569, A111.

Blanco-Cuaresma, S. *et al.* 2014b, volume 298 of IAU Symposium, 396.

Böcek Topcu, G. *et al.* 2015, MNRAS, 446, 3562.

Boesgaard, A. M., Lum, M. G., and Deliyannis, C. P. 2015, ApJ, 799, 202.

Boesgaard, A. M., Roper, B. W., and Lum, M. G. 2013, ApJ, 775, 58.

Boesgaard, A. M. and Tripicco, M. J. 1986, ApJL, 302, L49.

Booth, A. J. *et al.* 1984, MNRAS, 208, 147.

Bouchy, F. and Sophie Team 2006, in Tenth Anniversary of 51 Peg-b: Status of and prospects for hot Jupiter studies, ed. L. Arnold, F. Bouchy, and C. Moutou, 319.

Brandt, T. D. and Huang, C. X. 2015, ApJ, 807, 24.

Bresolin, F., Kennicutt, Jr., R. C., and Stetson, P. B. 1996, AJ, 112.

Bressan, A. *et al.* 2012, MNRAS, 427, 127.

Buurman, E. P. *et al.* 1986, A&A, 164, 224.

Cannon, A. J. and Pickering, E. C. 1901, Annals of Harvard College Observatory, 28, 129.

Cardon, B. L. *et al.* 1982, ApJ, 260, 395.

Carraro, G., Monaco, L., and Villanova, S. 2014, A&A, 568, A86.

Carrera, R. and Pancino, E. 2011, A&A, 535, A30.

Casagrande, L. *et al.* 2010, A&A, 512, A54.

Cass, S. 2018, IEEE Spectrum.

Castelli, F. and Kurucz, R. L. 2004, ArXiv Astrophysics e-prints.

Castro, M. *et al.* 2016, A&A, 590, A94.

Chaboyer, B., Green, E. M., and Liebert, J. 1999, AJ, 117, 1360.

Chambers, K. C. *et al.* 2016, ArXiv e-prints.

Chandrasekhar, S. 1939, Ciel et Terre, 55, 412.

Chen, L., Hou, J. L., and Wang, J. J. 2003, AJ, 125, 1397.

Claver, C. F. *et al.* 2001, ApJ, 563, 987.

Cohen, J. G. *et al.* 2003, ApJ, 588, 1082.

Culp, M. and Michailidis, G. 2007, in The 2007 Joint Statistical Meetings (JSM).

Daniel, S. *et al.* 1994, PASP, 106, 281.

Dekker, H. *et al.* 2000, in Optical and IR Telescope Instrumentation and Detectors, ed. M. Iye and A. F. Moorwood, volume 4008 of Proc. SPIE, 534.

Den Hartog, E. A. *et al.* 2011, ApJS, 194, 35.

Den Hartog, E. A. *et al.* 2014, ApJS, 215, 23.

Dias, W. S. *et al.* 2002, A&A, 389, 871.

Dobbie, P. D. *et al.* 2004, MNRAS, 355, L39.

Dominguez, I. *et al.* 1999, ApJ, 524, 226.

D'Orazi, V. and Randich, S. 2009, A&A, 501, 553.

Dotter, A. *et al.* 2008, ApJS, 178, 89.

Edvardsson, B. *et al.* 1993, A&A, 275, 101.

Elkan, C. and Noto, K. 2008, in Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '08, (New York, NY, USA: ACM), 213.

Flewelling, H. A. *et al.* 2016, ArXiv e-prints.

Francois, P. 1988, A&A, 195, 226.

Friel, E. D. 1995, ARA&A, 33, 381.

Friel, E. D. and Janes, K. A. 1993, A&A, 267, 75.

Friel, E. D. *et al.* 2002, AJ, 124, 2693.

Fry, A. M. and Carney, B. W. 1997, AJ, 113, 1073.

Fuhr, J. R., Martin, G. A., and Wiese, W. L. 1988, Journal of Physical and Chemical Reference Data, 17.

Fuhrmann, K., Axer, M., and Gehren, T. 1995, A&A, 301, 492.

Gaia Collaboration *et al.* 2018a, ArXiv e-prints.

Gaia Collaboration *et al.* 2018b, ArXiv e-prints.

Gaia Collaboration *et al.* 2016, A&A, 595, A1.

Galilei, G. 1610, Sidereus nuncius magna, longeque admirabilia spectacula pandens lunae facie, fixis innumeris, lacteo circulo, stellis nebulosis, ... Galileo Galileo : nuper a se reperti beneficio sunt observata in apprime vero in quatuor planetis circa Iovis stellam disparibus intervallis, atque periodis, celeritate mirabili circumvolutis ... atque Medicea sidera nuncupandos decrevit, (:).

Garz, T. 1973, A&A, 26, 471.

Geller, A. M., Latham, D. W., and Mathieu, R. D. 2015, AJ, 150, 97.

Giridhar, S. and Arellano Ferro, A. 1995, RMXAA, 31, 23.

Gray, D. F. and Johanson, H. L. 1991, PASP, 103, 439.

Habets, G. M. H. J. and Heintze, J. R. W. 1981, A&AS, 46, 193.

Hannaford, P. *et al.* 1982, ApJ, 261, 736.

Heinemann, K. 1926, Astronomische Nachrichten, 227, 193.

Heiter, U. *et al.* 2014, A&A, 561, A93.

Herschel, W. 1818, Philosophical Transactions of the Royal Society of London Series I, 108, 429.

Hobbs, L. M. and Pilachowski, C. 1986, ApJL, 309, L17.

Høg, E. *et al.* 2000, A&A, 355, L27.

Huang, Y. *et al.* 2015, MNRAS, 454, 2863.

Iben, I. 1991, APJ Suppl, 76, 55.

Iben, Jr., I. 1964, ApJ, 140, 1631.

Iben, Jr., I. 1967, ARA&A, 5, 571.

Iben, Jr., I. 1977, ApJ, 217, 788.

Johnson, H. L. and Knuckles, C. F. 1955, ApJ, 122, 209.

Kaeppeler, F. *et al.* 1990, ApJ, 354, 630.

Kalirai, J. S. and Richer, H. B. 2010, Philosophical Transactions of the Royal Society of London Series A, 368, 755.

Kalirai, J. S. *et al.* 2001, AJ, 122, 257.

Kaufer, A. *et al.* 1999, The Messenger, 95, 8.

Kennicutt, R. C. and Evans, N. J. 2012, ARA&A, 50, 531.

Kharchenko, N. V. *et al.* 2009, A&A, 504, 681.

Kharchenko, N. V. *et al.* 2016, A&A, 585, A101.

King, I. R. *et al.* 2005, AJ, 130, 626.

Kock, M. and Richter, J. 1968, ZA, 69, 180.

Kramida, A. *et al.* 2014a. NIST Atomic Spectra Database (ver. 5.2), [Online]. Available:`http://physics.nist.gov/asd` [2015, June 1].National Institute of Standards and Technology,Gaithersburg, MD.

Kramida, A. *et al.* 2014b. NIST Atomic Spectra Database (ver. 5.2), [Online]. Available:http://physics.nist.gov/asd [2015, June 1].National Institute of Standards and Technology,Gaithersburg, MD.

Kraus, A. L. and Hillenbrand, L. A. 2007, AJ, 134, 2340.

Kretz, W. C. 1900, Contributions from the Rutherford Observatory of Columbia University New York, 1.

Kroll, S. and Kock, M. 1987, A&AS, 67, 225.

Kuiper, G. P. 1938, ApJ, 88, 472.

Lambert, D. L. and Luck, R. E. 1978, MNRAS, 183, 79.

Lambert, D. L. and Warner, B. 1968, MNRAS, 138, 213.

Lawler, J. E. and Dakin, J. T. 1989, Journal of the Optical Society of America B Optical Physics, 6, 1457.

Lawler, J. E. *et al.* 2013, ApJS, 205(2), 11.

Lawler, J. E., Sneden, C., and Cowan, J. J. 2015, ApJS, 220(1), 13.

Lawler, J. E. *et al.* 2009, ApJS, 182, 51.

Lawler, J. E. *et al.* 2008, ApJS, 178(1), 71.

Lawler, J. E. *et al.* 2014, ApJS, 215, 20.

Leavitt, H. S. 1908, Annals of Harvard College Observatory, 60, 87.

Lind, K. *et al.* 2011, A&A, 528, A103.

Liu, F. *et al.* 2016, MNRAS, 457, 3934.

Loktin, A. V. 2000, Astronomy Letters, 26, 657.

Lum, M. G. and Boesgaard, A. M. 2018, ApJ.

Lyngå, G. 1983, Publications of the Astronomical Institute of the Czechoslovak Academy of Sciences, 56, 80.

Maderak, R. M. *et al.* 2013, AJ, 146, 143.

Marigo, P. *et al.* 2017, ApJ, 835, 77.

Martin, G. A., Fuhr, J. R., and Wiese, W. L. 1988, Atomic transition probabilities. Scandium through Manganese, volume 17, (New York, New York: American Institute of Physics (AIP) and American Chemical Society).

Maury, A. C. and Pickering, E. C. 1897, Annals of Harvard College Observatory, 28, 1.

Mayor, M. *et al.* 2003, The Messenger, 114, 20.

McWilliam, A. *et al.* 1995, AJ, 109, 2757.

McWilliam, A. and Rich, R. M. 1994, ApJS, 91, 749.

Meibom, S., Andersen, J., and Nordström, B. 2002, A&A, 386, 187.

Mermilliod, J.-C. and Mayor, M. 1999, A&A, 352, 479.

Messier, C. 1781. Catalogue des Nébuleuses et des Amas d'Étoiles (Catalog of Nebulae and Star Clusters). Technical report, Bureau des Longitudes.

Meylan, T. *et al.* 1993, ApJS, 85, 163.

Meynet, G., Mermilliod, J.-C., and Maeder, A. 1993, A&AS, 98, 477.

Missana, M. and Missana, N. 1995, AJ, 109, 1903.

Mocak, M. *et al.* 2009, A&A, 501, 659.

Moity, J. 1983, A&AS, 52, 37.

Montgomery, K., Janes, K., and Phelps, R. 1994, AJ, 108(2), 585.

Montgomery, K. A., Marschall, L. A., and Janes, K. A. 1993, AJ, 106, 181.

Moore, G. 1965, Electronics, 38.

Moultaka, J. *et al.* 2004, PASP, 116, 693.

Murray, N. 2011, ApJ, 729, 133.

Netopil, M. and Paunzen, E. 2013, A&A, 557, A10.

Netopil, M. *et al.* 2016, A&A, 585, A150.

Nissen, P. E. and Schuster, W. J. 1997, A&A, 326, 751.

Nordstroem, B., Andersen, J., and Andersen, M. I. 1997, A&A, 322, 460.

O'Brian, T. R. *et al.* 1991, Journal of the Optical Society of America B Optical Physics, 8, 1185.

Önehag, A. *et al.* 2011, A&A, 528, A85.

Pasquini, L., Randich, S., and Pallavicini, R. 2001, A&A, 374, 1017.

Paunzen, E. and Netopil, M. 2006, MNRAS, 371, 1641.

Percival, S. M., Salaris, M., and Kilkenny, D. 2003, A&A, 400, 541.

Perryman, M. A. C. *et al.* 1998, A&A, 331, 81.

Perryman, M. A. C. *et al.* 1997, A&A, 323, L49.

Pilachowski, C., Saha, A., and Hobbs, L. M. 1988, PASP, 100, 474.

Pilachowski, C. A. and Hobbs, L. M. 1988, PASP, 100, 336.

Platais, I. 1991, A&A Suppl, 87, 69.

Prochaska, J. X. *et al.* 2000, AJ, 120, 2513.

Prugniel, P. and Soubiran, C. 2001, A&A, 369, 1048.

Rafelski, M. and Zaritsky, D. 2005, AJ, 129, 2701.

Ramírez, I. and Allende Prieto, C. 2011, ApJ, 743, 135.

Ramírez, I. and Meléndez, J. 2005, ApJ, 626, 465.

Reddy, A. B. S., Giridhar, S., and Lambert, D. L. 2012, MNRAS, 419, 1350.

Reddy, A. B. S., Giridhar, S., and Lambert, D. L. 2013, MNRAS, 431, 3338.

Reino, S. *et al.* 2018, MNRAS, 477, 3197.

Rentzsch-Holm, I. 1996, A&A, 312, 966.

Rieke, G. H. and Lebofsky, M. J. 1985, ApJ, 288, 618.

Roberts, I. 1893, A Selection of Photographs of Stars, Star-Clusters and Nebulae, together with Information concerning the Instruments and the Methods employed in the pursuit of Celestial Photography, (London, England: The Universal Press).

Rohlfs, K. and Vanysek, V. 1962, Astronomische Abhandlungen der Hamburger Sternwarte, 5, 341.

Russeil, D. 2003, A&A, 397, 133.

Russell, H. N. 1914, Popular Astronomy, 22, 275.

Ryabchikova, T. *et al.* 2015, PhyS, 90(5).

Salaris, M. *et al.* 2015, A&A, 583, A87.

Salaris, M., Weiss, A., and Percival, S. M. 2004, A&A, 414, 163.

Sandage, A. 1955, Leaflet of the Astronomical Society of the Pacific, 7, 57.

Savanov, I. S., Huovelin, J., and Tuominen, I. 1990, A&AS, 86, 531.

Schlafly, E. F. and Finkbeiner, D. P. 2011, ApJ, 737, 103.

Schlegel, D. J., Finkbeiner, D. P., and Davis, M. 1998, ApJ, 500, 525.

Schnabel, R., Kock, M., and Holweger, H. 1999, A&A, 342, 610.

Scott, D. 1992, Multivariate Density Estimation: Theory, Practice, and Visualization, (Chichester, NY: John Wiley & Sons.).

Sestito, P., Randich, S., and Pallavicini, R. 2004, A&A, 426, 809.

Shapley, H. 1917, ApJ, 45.

Shapley, H. and Richmond, M. L. 1921, ApJ, 54.

Shu, F. H., Adams, F. C., and Lizano, S. 1987, ARA&A, 25, 23.

Skrutskie, M. F. *et al.* 2006, AJ, 131, 1163.

Smith, G. and Raggett, D. S. J. 1981, Journal of Physics B Atomic Molecular Physics, 14, 4015.

Sneden, C. 1973, ApJ, 184, 839.

Sneden, C. *et al.* 2012. Astrophysics Source Code Library.

Sneden, C. *et al.* 2014, ApJS, 214, 26.

Sobeck, J. S., Lawler, J. E., and Sneden, C. 2007, ApJ, 667(2), 1267.

Stephens, A. 1999, AJ, 117, 1771.

Tabernero, H. M., Montes, D., and González Hernández, J. I. 2012, A&A, 547, A13.

Takeda, Y. 2003, A&A, 402, 343.

Takeda, Y., Ohkubo, M., and Sadakane, K. 2002, PASJ, 54, 451.

Tofflemire, B. M. *et al.* 2014, AJ, 148, 61.

van Rossum, G., Warsaw, B., and Coghlan, N. 2013.

Ventura, P., Carini, R., and D'Antona, F. 2011, MNRAS, 415, 3865.

Vereshchagin, S. V. and Chupina\inst, N. V. 2013, Astronomische Nachrichten, 334, 892.

Vogt, S. S. *et al.* 1994, in Instrumentation in Astronomy VIII, ed. D. L. Crawford and E. R. Craine, volume 2198 of Proc. SPIE, 362.

Wallace, L. *et al.* 2011, ApJS, 195, 6.

Wells, D. C., Greisen, E. W., and Harten, R. H. 1981, A&AS, 44, 363.

Wenger, M. *et al.* 2000, A&AS, 143, 9.

Werchan, F. and Zaritsky, D. 2011, AJ, 142, 48.

Wickliffe, M. E. and Lawler, J. E. 1997, ApJS, 110, 163.

Wiese, W. L. and Martin, G. A. 1980, Wavelengths and transition probabilities for atoms and atomic ions: Part 2. Transition probabilities, NSRDS-NBS Vol. 68., (Washington D.C.: NSRDS-NBS, US Department of Commerce, National Bureau of Standards).

Wood, M. P. *et al.* 2013, The Astrophysical Journal Supplement Series, 208(2), 27.

Wood, M. P. *et al.* 2014, The Astrophysical Journal Supplement Series, 211(2), 20.

Worthey, G. and Jowett, K. J. 2003, PASP, 115, 96.

Yang, X. L., Chen, Y. Q., and Zhao, G. 2015, AJ, 150, 158.

Yong, D., Carney, B. W., and Friel, E. D. 2012, AJ, 144, 95.

Zhu, X. 2008. Semi-supervised learning literature survey. Technical Report 1530, Computer Sciences, University of Wisconsin-Madison.